

UNIVERSIDAD COMPLUTENSE DE MADRID  
FACULTAD DE CIENCIAS DE LA INFORMACIÓN



UNIVERSIDAD  
**COMPLUTENSE**  
MADRID

# LA NUBE COMO APOYO A LA PRODUCCIÓN AUDIOVISUAL

TFM2 – MANUEL MONTES – MANMON01@UCM.ES

**DIRIGIDO POR:**  
ALBERTO GARCÍA GARCÍA  
JUAN GARCÍA CREGO

**MÁSTER EN COMUNICACIÓN AUDIOVISUAL PARA  
LA ERA DIGITAL**  
DEPARTAMENTO DE COMUNICACIÓN AUDIOVISUAL Y PUBLICIDAD I

JUNIO DE 2016

# Índice

1	Introducción .....	1
1.1	Amazon y su modelo de negocio .....	2
2	Marco teórico .....	5
3	Hipótesis.....	8
4	Objetivos .....	9
5	Metodología .....	9
6	La nube: definición, modelos y servicios.....	11
6.1	Beneficios .....	13
6.2	Virtualización.....	14
6.3	Centros de datos .....	15
6.3.1	Economías de alcance.....	16
6.3.2	Regiones y zonas de disponibilidad .....	18
7	Amazon Web Services .....	18
7.1	Amazon Ec2.....	18
7.2	Cloudfront.....	21
7.3	Amazon VPC .....	22
7.3.1	Direct Connect.....	24
7.4	Pago por uso.....	24
7.5	AWS y renderfarms .....	25
7.6	Escalado y autoescalado .....	25
7.7	Otros servicios para considerar.....	26
8	Perspectivas de mercado .....	28
8.1	Indicaciones previas.....	28
8.1.1	Tarifas de Amazon Web Services .....	29
8.2	Caso 1: computación.....	29
8.2.1	Situación .....	29
8.2.2	Infraestructura.....	29
8.2.3	Costes .....	31
8.2.4	Otros usos .....	32

8.2.5	Reducción de costes .....	33
8.3	Caso 2: web interactiva.....	35
8.3.1	Situación .....	35
8.3.2	Infraestructura.....	35
8.3.3	Costes .....	36
8.3.4	Reducción de costes .....	39
8.3.5	Geocacheado.....	39
8.4	Resumen.....	39
8.5	Posibles aplicaciones alternativas .....	40
9	Conclusiones .....	41
10	Recursos audiovisuales.....	43
11	Bibliografía.....	44
12	Citas de internet .....	44

# Resumen

En esta investigación analizaremos las capacidades de la nube, centrándonos en la computación (entendiendo esta como la capacidad de procesamiento) y el almacenamiento de datos geolocalizados (redes de entrega de contenidos). Estas capacidades, unidas al modelo de negocio que implica un coste cero por aprovisionamiento y el pago por uso, además de la reducción de costes mediante sinergias y optimización de los centros de datos, suponen un escenario interesante para las PYMES dedicadas a la producción audiovisual. Utilizando los servicios en la nube, una empresa pequeña puede, en cuestión de minutos, alcanzar la potencia de procesamiento y distribución de gigantes de la producción audiovisual, sin necesidad de ocupar sus instalaciones ni pagar por adelantado el alquiler de los equipos. Describiremos los servicios de AWS (Amazon Web Services) que serían útiles a una empresa, el uso que darían de dichos servicios y su coste, y lo compararemos con el presupuesto por realizar una misma instalación físicamente en dicha empresa, teniendo que comprar e instalar los equipos.

Palabras clave:

Amazon Web Services, nube, computación, producción audiovisual, pymes.

# 1 INTRODUCCIÓN

---

La nube está en boca de todos, más aún cuando las últimas noticias la relacionan con filtraciones (un asunto, el de la seguridad, que también trataremos en este documento), pero no es una novedad tecnológica. La capacidad de trabajar en la nube existe desde la invención de la informática, ya que, en esencia, se trata de esos mismos sistemas informáticos deslocalizados.

Amazon comenzó a expandir sus centros de datos para soportar la carga que suponía mantener su página web de comercio online. Posteriormente comenzó a vender parte de esa capacidad a sus socios y en 2006 abrió el servicio de Amazon Web Services (o AWS) al público general (Amazon, s.f.). Existen otras compañías que ofrecen servicios en la nube similares, como Google o Microsoft, pero lo particular de AWS (y por ello nos centraremos en este servicio) es que su interfaz se ha hecho tan intuitiva que no requiere conocimientos informáticos para comenzar a funcionar. Mediante dicha interfaz de usuario se puede empezar a operar en varios clics. Existe incluso un mercado de máquinas virtuales llamado AWS Marketplace en el que, en unos minutos, se pueden poner en funcionamiento instancias de AWS con aplicaciones ya preinstaladas: gestores de contenido, herramientas de análisis y monitorización, servidores de copia de seguridad, etc.

En España, las PYME suponen el 99% del total de empresas, 3.110.522 (Dirección General de Industria y de la PYME, 2015). A pesar de ser mayoría, concentran un nivel de ingresos moderado: el 97% ingresa menos de 2 millones de euros. Las empresas cuyo mercado es la producción audiovisual necesitan una infraestructura informática que soporte las cargas de trabajo: renderizado de vídeo, efectos y 3D, conversión de formatos, y además recientemente, muchas empresas también realizan audiovisual para la red, lo que incluye en ocasiones la necesidad de disponer de servidores que gestionen un nivel elevado de visitas de usuarios, y además picos puntuales en determinadas campañas.

Una empresa de este tipo, con unas ganancias tan reducidas y un mercado de trabajo (el audiovisual español) que no es constante ni previsible, no puede permitirse invertir en una infraestructura informática que le permita procesar más rápidamente, con mejor calidad y de forma paralela. Por tanto, una PYME audiovisual queda relegada a trabajos poco importantes, con facturas pequeñas y baja carga de trabajo (la que puede soportar su equipamiento). Además, cuando se trata de procesamiento de efectos y 3D, los tiempos muertos hacen el trabajo mucho menos eficiente.

Gracias a AWS, cualquier empresa podría, en cuestión de minutos, crear una infraestructura desechable, por la que solo pagaría mientras utiliza y que podría eliminar después de forma completa o guardar una “referencia” a un coste muy reducido, con la que, en unos minutos, dicha infraestructura podría volver a estar funcionando en el mismo estado en que se dejó.

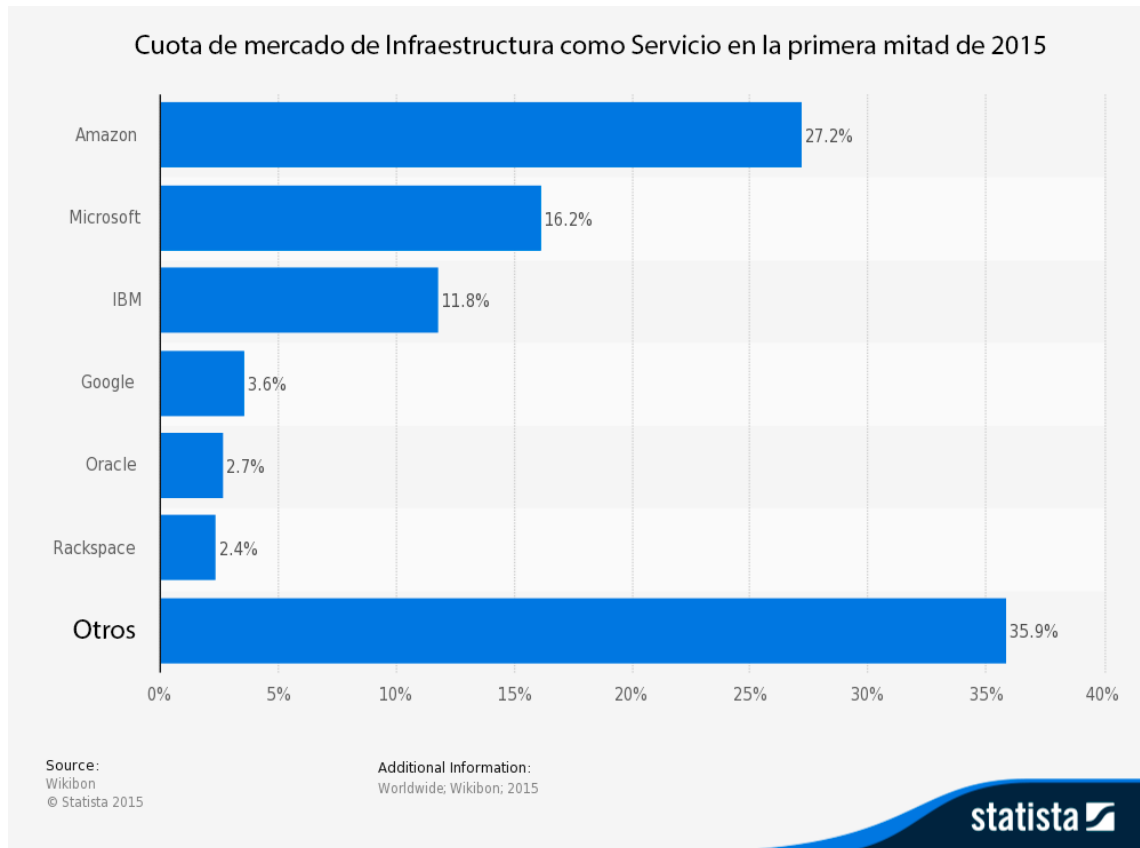
Además, al ser una infraestructura externa, se eliminan los tiempos muertos que suponen los procesos mencionados anteriormente.

Al tratarse de centros de datos situados en localizaciones concretas, la transferencia de datos supone un coste y también toma un tiempo. En este documento, por simplificar, no eliminaremos el coste de esta transferencia, pero daremos por hecho que las capacidades de la red son suficientes como para no suponer un aumento significativo del tiempo. Es cierto que las redes que se venden al público aún tienen sus limitaciones, pero intuimos que es una cuestión más bien comercial, dado que hay empresas que han demostrado que la capacidad da para más (Cid, 2014).

## 1.1 AMAZON Y SU MODELO DE NEGOCIO

Amazon se ha convertido en uno de los proveedores de servicio en la nube con mayor éxito. Según datos de octubre de 2009 tan solo uno de sus componentes, EC2, ya generaba más de 220 millones de dólares (Sosinsky, 2012, pág. 245). Además, se calculaba que el negocio incrementaría sus resultados en un 10 - 20% anual. Según los últimos datos, los ingresos de Amazon Web Services ascienden a 2,57 mil millones de dólares (Lynley, 2016).

*Ilustración 1: Cuota de mercado de IaaS en la nube. Traducción de la original libre*



Muchas de las empresas que están teniendo éxito en internet confían su infraestructura en la nube a AWS (Amazon case Studies). Netflix, Adobe, Pinterest, Zynga, Lionsgate, Comcast e incluso las españolas Securitas Direct y F.C. Barcelona están en su lista de clientes. Se calcula que para 2017 los ingresos de AWS superen los 11 mil millones de dólares (Greene, 2015).

Aparte de la aparente ventaja económica de Amazon respecto a otros proveedores, existen otros motivos para haber seleccionado esta plataforma y no cualquier otra como base de análisis en esta investigación. Haremos a continuación una breve comparación con Microsoft, por ser el inmediatamente posterior en el gráfico (ilustración 1); y con Google, por ser de los más conocidos.

## **CDN**

Red de Entrega de Contenidos o CDN (por sus siglas en inglés) hace referencia a un sistema mediante el cual un archivo o archivos se replican en servidores físicos distribuidos geográficamente (ver ampliación en epígrafe 7.2).

En el caso de Google, este servicio ni siquiera existe de manera separada. La red de entrega de contenidos de Google forma parte de su balanceador de carga (Google CDN, s.f.). Esto implica que no puede ser usado con orígenes externos (contenido no alojado en las propias máquinas virtuales que ofrece el proveedor), algo que sí permiten tanto Microsoft como Amazon. En el caso de Microsoft (Azure CDN, s.f.), aunque sí está permitido el uso de la red para distribuir contenido externo, aún no es posible hacerlo mediante protocolos de tráfico seguro SSL. Está previsto que esta funcionalidad se incluya en 2016, pero ya se ha retrasado en varias ocasiones.

Teniendo en cuenta estas características, podemos decir que Amazon es la plataforma que ofrece la mayor cantidad de ellas.

## **Máquinas virtuales**

Las máquinas virtuales son sistemas que reúnen una serie de características predefinidas (ver epígrafe “Amazon Ec2”) en cuanto a procesadores, memoria, acceso a la red, etc.

La principal diferencia entre los tres servicios la encontramos en las instancias puntuales (ver epígrafe “economías de alcance”). Microsoft de entrada no ofrece este servicio (ni está previsto a corto plazo). Google maneja este tipo de instancias de forma diferente. En dicho servicio son conocidas como “VM prioritaria” (Google Cloud Preemptible VM, s.f.) y están configuradas para ser liquidadas (*terminated*) cuando el sistema lo requiere, es decir, que los datos que almacenan se perderían. En el caso de Amazon, la instancia o máquina virtual se

detiene pero sus datos no se eliminan. Además, este tipo de modelo supone una reducción considerable en el precio que depende de la disponibilidad de recursos sin usar.

De los tres proveedores, Amazon es el único que ofrece, de momento, instancias con procesadores gráficos (NVIDIA). Está previsto que Microsoft lo haga durante 2016, incorporando máquinas con tarjetas gráficas NVIDIA Tesla. No obstante, la fecha estaba fijada en otoño y aún no se sabe nada oficialmente (Staveley, s.f.). Este tipo de instancias son importantes a la hora de procesar gráficos 3D e imagen en tiempo real, dado que esta tarea es especialidad de las tarjetas gráficas, que están más optimizadas para ello que las CPU.

Por último, cada uno de los proveedores ofrece una serie de instancias predefinidas, de las cuales los clientes pueden hacer uso de una o varias del mismo tipo. La cantidad, a la finalización de esta investigación, es la siguiente:

- Amazon: 55
- Google: 18
- Microsoft: 45

Aunque Google permite la creación de máquinas virtuales personalizadas, donde el usuario selecciona la cantidad de recursos que desea, esta herramienta permite disponer de un máximo de 32 núcleos de CPU y 208 GB de memoria, muy alejado de lo que es posible con las máquinas virtuales predefinidas de Amazon.

### **Reducción de coste por uso sostenido**

Al utilizar las instancias proporcionadas por Google, a partir de ciertos niveles de uso se obtiene un descuento por uso sostenido (Google Compute Engine, s.f.), que se aplica automáticamente y supone pagar desde el 100% (con un uso inferior al 25% mensual) de la tarifa hasta el 40% (si el uso es del 100% mensual). Según la compañía, esto sería equivalente a un descuento efectivo del 30% como máximo.

En el caso de Amazon existe la reserva de instancias, mediante la cual, gracias a un compromiso de uso por un tiempo determinado, se obtiene una rebaja en el precio. Este descuento no se obtiene automáticamente como ocurre en Google, sino que el usuario debe solicitarlo previamente. Además, una vez solicitado ese compromiso, no puede anularse (aunque existe la posibilidad de reventa), mientras que la reducción de costes en Google no implica compromiso alguno. A cambio de este compromiso, la rebaja que puede obtenerse en el precio llega hasta el 75% del precio habitual (Amazon EC2 reserved instances, s.f.).

Para sintetizar, estos son algunos de los puntos diferenciales de Amazon frente a otros proveedores:



- Ofrece una red de entrega de contenidos (CDN) con la posibilidad de distribuir orígenes externos mediante conexión segura (SSL), mientras Microsoft no permite la conexión segura y Google ni siquiera ofrece esta posibilidad para contenidos externos.
- La cantidad de máquinas virtuales predefinidas ofrecida por Amazon es mayor, lo que permite flexibilidad a la hora de elegir qué tipos (o combinación de tipos) serán más adecuados para distintos proyectos. Además, es el único de los tres proveedores que ofrece instancias con gráficas dedicadas.
- Respecto a la reducción de costes, las máquinas virtuales puntuales conservan los datos una vez detenidas, mientras en Google los datos se eliminan y Microsoft ni siquiera permite esta posibilidad. Además, cuando se utilizan máquinas virtuales a largo plazo, Amazon ofrece un mayor descuento a cambio de un compromiso de uso.

Esta serie de condiciones, algunas de las cuales se desarrollarán más adelante, indican que Amazon es, en la actualidad, el proveedor de servicios con mayor flexibilidad en cuanto a precio y opciones, y por ello desarrollaremos la investigación con sus servicios. Esto no impide que, en un futuro, otros proveedores puedan volverse más adecuados si aumentan su oferta en cuanto a las características o reducen precios.

## 2 MARCO TEÓRICO

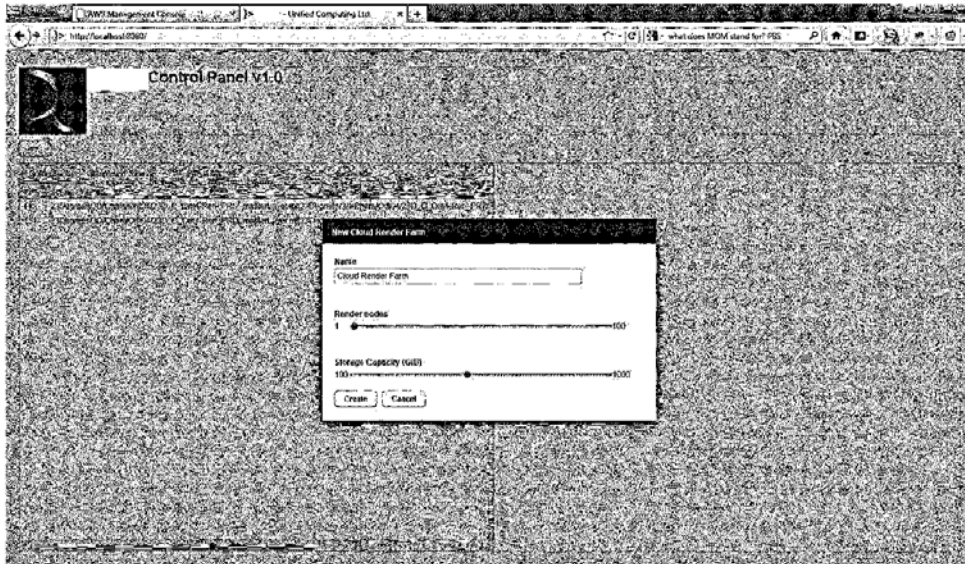
---

La relación entre la comunicación audiovisual y la nube no dispone de mucha literatura. En la mayoría de los casos, el desarrollo de tecnologías en la nube para empresas de comunicación requiere los servicios de personal informático especializado.

No hay que diversificar demasiado ni ir muy atrás en el tiempo para encontrar el primer intento de desarrollar servicios en la nube orientados al renderizado. En una patente solicitada en 2014 (Estados Unidos Patente nº US20140237373 A1, 2014) se aprecia un sistema diseñado para el renderizado de imágenes 3D en la nube. Dicha patente pretende proveer de un sistema fácil e intuitivo para solicitar máquinas virtuales en la nube, diseñado para personas no especializadas en informática.

Mediante una interfaz gráfica, este sistema permitiría seleccionar el número de nodos para renderizar (lo que se traduciría en cierto número de máquinas virtuales trabajando a la vez) y el espacio en el disco duro necesario.

*Ilustración 2: panel de control diseñado en la patente*



“De esta forma, las granjas de render en la nube pueden ser habilitadas fácilmente con mínima dificultad y habilidades. Una vez creadas, el renderizado puede ser solicitado arrastrando el trabajo hasta un icono en la interfaz de usuario.” (Traducción del autor)

Esta patente tiene como base una infraestructura presentada por la empresa Axceleon en 2010. Su sistema CloudFuzion (Axceleon Inc., 2010), a través de una interfaz de usuario (llamada CloudFuzion Desktop) permite solicitar y monitorizar trabajos de renderizado.

Ambos sistemas tienen en común la utilización de los servicios proporcionados por Amazon para el aprovisionamiento de las máquinas virtuales<sup>1</sup>.

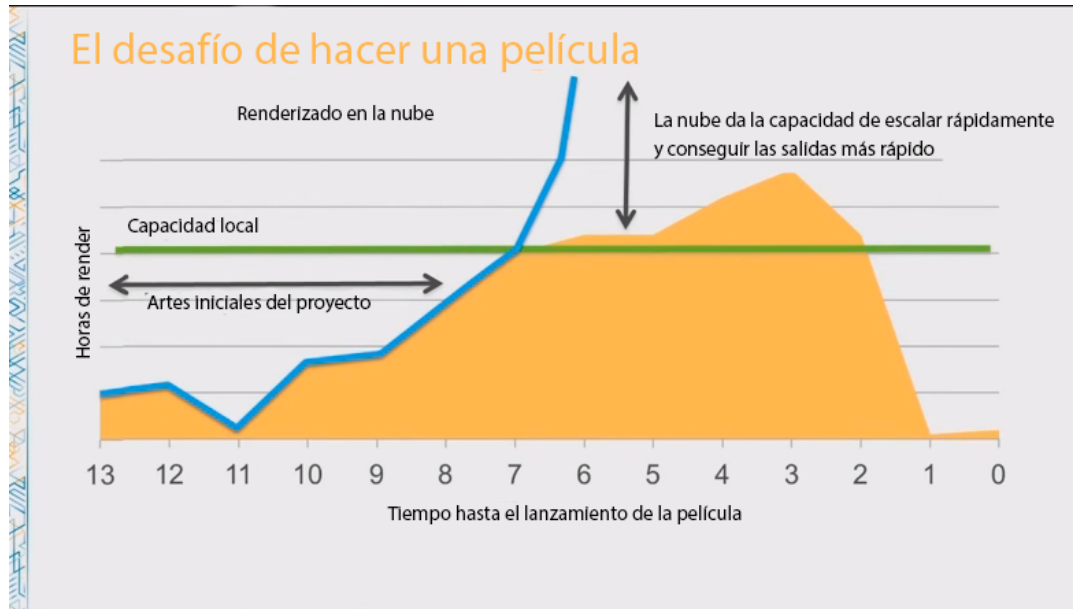
Encontramos otra referencia al uso de la computación en la nube para renderizado en una conferencia de Amazon reciente, de 2015. En esta conferencia (AWS re:Invent, 2015) se presentaban las ventajas del render en la nube con respecto al sistema tradicional en granjas o equipos locales.

---

<sup>1</sup> Las máquinas virtuales en el entorno de Amazon son denominadas “instancias” y hacen referencia a una serie de características: núcleos, memoria, etc.

Además, la conferencia confronta el uso hipotético de estas herramientas por parte de dos empresas: un pequeño estudio y la productora Disney. Esta presentación tiene mucho que ver con el objetivo de este escrito.

Ilustración 3: escalabilidad de renderizado en la nube frente a local (AWS re:invent). Traducción de la original libre



Las necesidades de estas dos empresas son en esencia las mismas: capacidad para utilizar miles de núcleos de procesamiento sin ninguna inversión previa y la creación de una estructura “desechable” basada en proyectos concretos.

En una de las pruebas realizadas por la compañía, llegaron a tener hasta un total de 40.000 núcleos funcionando simultáneamente en un mismo proyecto. Si bien no se especifica en la conferencia nada sobre estos núcleos, es lógico pensar que se trataba de procesadores Intel Xeon E5-2666 v3<sup>1</sup>, que son los utilizados en instancias optimizadas para computación.

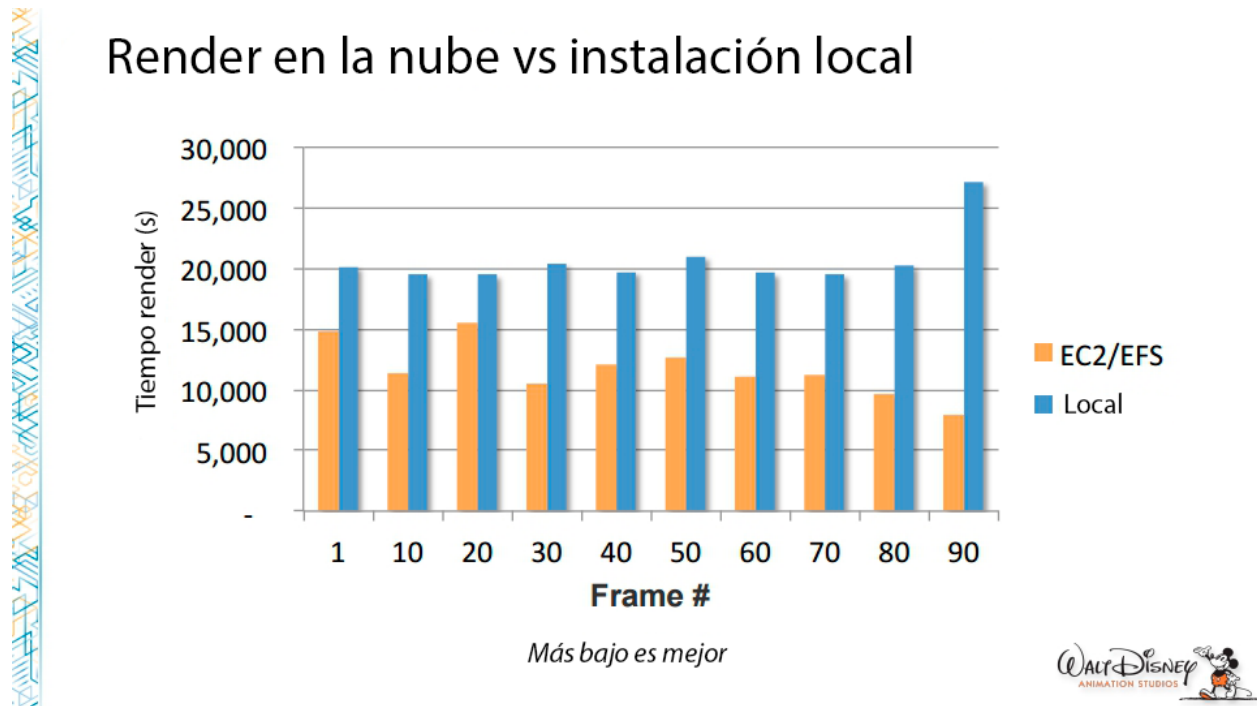
Además de la escalabilidad inmediata y sin coste de instalación, el renderizado en la nube muestra ser más eficiente que el realizado en equipos locales (*on-premise*).

Existe muy poca literatura científica sobre este tipo de uso concreto de la nube. En este documento iremos un poco más allá de lo ya escrito, proponiendo modelos y estructuras concretas.

---

<sup>1</sup> Cada procesador de 64 bits cuenta con 10 núcleos a 2.6 – 3.3 GHZ y es capaz de soportar hasta 768 GB de memoria RAM.

Ilustración 4: tiempo tomado por cada frame renderizado en la nube y en local. Traducción de la original libre



### 3 HIPÓTESIS

Una empresa audiovisual cuya actividad sea la creación de contenido necesita principalmente dos pilares para funcionar: la creatividad y la potencia informática. El primero es un factor humano que, si bien no es fácil de conseguir, tampoco es exclusividad de las grandes empresas. Una PYME puede contar con los mejores creativos del sector, y sin embargo verse limitada por la potencia informática. Esto es así porque el procesamiento de material (ya sea vídeo, efectos o imágenes 3D) requiere de una potencia que aumenta con la complejidad de cada trabajo, algo que no ocurre con la creatividad. Además, el material informático tiene un coste elevado, y la necesidad de adquirirlo para trabajos puntuales podría no ser rentable para la empresa.

Ante esta perspectiva, la potencia de computación y la capacidad de distribución de contenido disponible en la nube puede ser una solución (puntual, como ya indicamos) para eliminar la limitación informática.

Teniendo en cuenta este punto de partida, planteamos la siguiente hipótesis:

“El alquiler de servicios en la nube facilita a las PYMES eliminar la limitación tecnológica que, debido al elevado coste de instalación, les impide asumir producciones de mayor nivel técnico”.

## 4 OBJETIVOS

---

Teniendo en cuenta los antecedentes teóricos de la cuestión y las pretensiones de esta investigación, se pretende demostrar o desmentir las siguientes afirmaciones:

- Analizar la creación de entornos virtuales de renderizado en Amazon Web Services y el nivel de conocimiento necesario.
- Desarrollar la creación de infraestructuras desechables y analizar el coste repercutido en la empresa.
- Analizar y comparar el coste de una infraestructura local y su mantenimiento con una infraestructura en la nube de Amazon.
- Describir la función de escalado y sus particularidades en la nube, así como sus características.
- Describir las redes de entrega de contenidos y sus capacidades de caché y distribución geolocalizada.
- Analizar la seguridad de la nube y de las comunicaciones entre la nube y las instalaciones de los usuarios.
- Proponer un esquema de modelo de negocio que aproveche las características de la nube.

La nube y sus tecnologías se analizan frecuentemente desde el punto de vista de internet (redes sociales, páginas web, aplicaciones) y el análisis de datos (*big data*), de manera que se pierde la perspectiva de que en esencia son ordenadores gigantes trabajando en centros optimizados. La preparación de las instalaciones, además del material informático en sí, requiere una inversión inicial inasumible para una empresa cuya producción es variable. Desde este punto de vista, disponer de un proveedor de “alquiler” de potencial informático puede suponer que se elimine la barrera de entrada a una franja del mercado con necesidades más altas.

## 5 METODOLOGÍA

---

Por las características que hemos descrito en el marco teórico, nos encontramos ante un estudio de tipo exploratorio (Hernández Sampieri , Fernández Collado, y Baptista Lucio (1991), pág. 70):

“Los estudios exploratorios se efectúan, normalmente, cuando el objetivo es examinar un tema o problema de investigación poco estudiado o que no ha sido abordado antes. Es decir, cuando la revisión de la literatura reveló que únicamente hay guías no investigadas e ideas vagamente relacionadas con el problema de estudio. [...] Los estudios exploratorios nos sirven para aumentar el grado de familiaridad con

fenómenos relativamente desconocidos, obtener información sobre la posibilidad de llevar a cabo una investigación más completa sobre un contexto particular de la vida real...”

A lo largo del documento describiremos los servicios del ecosistema de Amazon Web Services que serían útiles a una empresa de producción audiovisual a la hora de acelerar el proceso de renderizado y distribución de contenidos. Para realizar dicha descripción inicial acudiremos a varias fuentes, principalmente la documentación técnica disponible en AWS y varias publicaciones que han descrito la nube. La intención de esta primera parte es acercar el funcionamiento de este servicio a los usuarios generales, que no necesariamente dispondrán de conocimientos informáticos. En una PYME, por ejemplo, no es habitual que exista un técnico informático permanente en la plantilla, dado que su utilidad se reduce a soluciones puntuales.

Junto con la descripción de los servicios y basándonos en las tarifas oficiales del centro de datos de la región de Irlanda (este concepto se explicará más adelante) se presentarán breves ejemplos de coste de almacenamiento, procesamiento y distribución.

Para facilitar la comprensión del entorno de AWS y su funcionamiento, realizaremos pruebas de instalación y rendimiento que serán grabadas e incluidas con el documento. Esto permitirá observar y reproducir el funcionamiento de las instancias de computación.

Las pruebas de rendimiento se realizarán con el software Cinebench (maxon.net, s.f.) que lleva a cabo el procesado de una imagen y devuelve el resultado en forma de puntuación. Esta imagen contiene 2.000 objetos, 300.000 polígonos y además luces, sombras, desenfoques y otros efectos que aumentan el tiempo de procesado.

“El entorno de pruebas usa todos los procesadores para crear un escenario 3D fotorealístico. Esta escena usa algoritmos para forzar todos los núcleos disponibles. De hecho, CINEBENCH puede probar sistemas con hasta 256 hilos de procesador. Esta escena contiene 2.000 objetos, que suponen más de 300.000 polígonos, y usa reflejos finos y difusos, luces de área, sombras, sombreado procesal, antialias y mucho más. El resultado se muestra en puntos. Cuanto más alto, más rápido es el procesador.” (Traducción del autor)



Además, se utilizará también Realbench, un software que permite realizar test de rendimiento y estrés mediante el uso de herramientas *open source*<sup>1</sup> como Blender<sup>2</sup>, GIMP<sup>3</sup>; etc. Realbench ya incluye una prueba de codificación h.264, pero igualmente y con el objetivo de probar el rendimiento del reciente códec h.265, se utilizarán otras dos herramientas: Adobe Media Encoder (en su versión CC) y HandBrake, un software *open source* de codificación de archivos. Todas estas herramientas aportan una calificación objetiva<sup>4</sup> que permitiría a cualquier usuario o empresa comparar los resultados con los de su propia instalación y calcular la diferencia teórica en términos de tiempo y coste.

Finalizaremos con un estudio de las perspectivas de mercado. En dicho estudio plantearemos dos casos de uso, mediante los que se propondrá un encargo a una teórica empresa audiovisual y se plantearán las estructuras utilizadas dentro de la nube y su conexión con los sistemas físicos de la compañía. Además, se desglosarán los costes de dicha instalación realizada en la nube y en el emplazamiento (*on-premise*) y se compararán los costes. Ambas posibilidades son muy distintas, ya que, en un caso, se trataría de la creación de una estructura desechable que solo se utilizará para trabajos concretos y se pagará en función del tiempo de uso. Contrariamente, una instalación física requiere de un desembolso económico inicial que no es retornable, es decir, una vez instalada la infraestructura, debe conservarse en la ubicación. No obstante, esta comparación nos permitirá averiguar si existen ventajas en la elección de la nube como sistema para cargas de trabajo puntuales o si, opuestamente, utilizar la nube supondría un coste demasiado elevado como para resultar rentable.

## 6 LA NUBE: DEFINICIÓN, MODELOS Y SERVICIOS

---

El uso de la palabra “nube” para definir a estas tecnologías tiene ciertas implicaciones. Esencialmente, la nube hace referencia a dos conceptos (Sosinsky, 2012):

- **Abstracción:** Las tareas se ejecutan en sistemas físicos que no están identificados (más allá de referencias). Además, el acceso a estas tareas es posible desde cualquier lugar y dispositivo.

---

<sup>1</sup> El software *open source* (o de código abierto) es distribuido y desarrollado libremente. Cualquier usuario puede tener acceso al código, modificarlo y crear productos derivados.

<sup>2</sup> Software *open source* de modelado 3D

<sup>3</sup> Software *open source* de tratamiento de imagen.

<sup>4</sup> Algunas como un número entero (CineBench) y otras expresada en tiempo.

- Virtualización: los recursos se reúnen y se comparten desde un fondo común (los centros de datos). Existe la posibilidad de multitenencia<sup>1</sup> y la escalabilidad es sencilla y ágil.

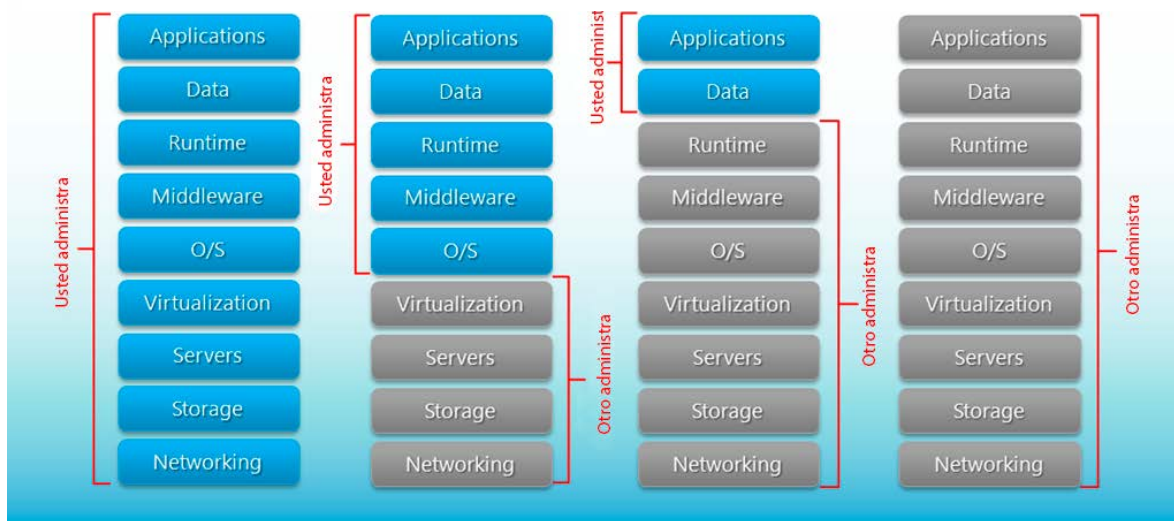
Según la definición del NIST (*National Institute of Standards and technology*) existen cuatro modelos de implementación, según el propósito y la naturaleza de la ubicación: nube pública, privada, híbrida y comunitaria.

La nube pública es, brevemente, aquella en la que un proveedor dispone de los recursos (equipos, etc.) y los pone a disposición de organizaciones y particulares, que solo pagan por el uso (coste operacional). La privada, en cambio, es aquella diseñada específicamente para una organización, con lo cual esta asume el coste de instalación además del operacional (es el caso más parecido a una instalación *on-premise*). La nube híbrida es una combinación de ambas.

Los modelos de servicio ofrecidos dependen del nivel de profundidad que alcance el usuario, es decir, cuánto se adentra en los recursos proporcionados. Ordenados de mayor a menor profundidad, existen tres modelos de servicio: IaaS (infraestructura como servicio), PaaS (plataforma como servicio) y SaaS (software como servicio).

## División de responsabilidades

Ilustración 5: Cloud Computing: IaaS, PaaS, SaaS (<http://programacion.jias.es/2015/05/cloud-computing-iaas-paas-saas/>). Traducción de la original libre



<sup>1</sup> Capacidad de un recurso informático de ser utilizado por distintos consumidores. Es la base de la eficacia y la rentabilidad de la nube. Mediante la virtualización, un mismo procesador o disco duro puede servir a varios usuarios a la vez.



(plataforma como servicio) y SaaS (Software como servicio). Nos quedaremos solo con la definición del primer modelo ya que es el que nos interesa, el resto no proporciona una profundidad suficiente para nuestro cometido.

La infraestructura como servicio proporciona al usuario máquinas, almacenamiento y demás hardware disponible. El proveedor es responsable de gestionar todos estos recursos, y el usuario de todo lo demás: sistema operativo, software, etc. IaaS proporciona máquinas virtuales “vacías” con un sistema operativo limpio que el usuario puede escoger. Amazon EC2 (*Elastic Compute Cloud*) es un claro ejemplo de IaaS, y es el servicio que describiremos y utilizaremos en este documento.

Amazon (Amazon EC2, s.f.) es responsable (por contrato) de proporcionar un cierto nivel de fiabilidad (medido en tiempo de ejecución de recursos).

“Amazon EC2 ofrece un entorno de elevada confianza en el que las instancias de sustitución se pueden enviar con rapidez y anticipación. El servicio se ejecuta en los centros de datos y la infraestructura de red acreditados de Amazon. El compromiso del Acuerdo de nivel de servicios de Amazon EC2 es ofrecer una disponibilidad del 99,95% en todas y cada una de las regiones de Amazon EC2.”

## 6.1 BENEFICIOS

Según (Sosinsky, 2012, pág. 43) la definición del NIST también indica que la nube debe poseer ciertas características esenciales:

- Autoservicio bajo demanda: el cliente debe poder disponer de los recursos que seleccione sin que tenga que mediar ningún agente.
- Acceso amplio desde la red: los recursos deben ser accesibles desde la red utilizando métodos<sup>1</sup> estándar.
- Fondo de recursos: Los recursos físicos se virtualizan y reasignan según sea necesario, de manera que existe compatibilidad con el uso en multitenencia y además la ubicación física de los recursos permanece oculta.
- Elasticidad rápida: Los recursos pueden suministrarse de forma rápida y elástica (sic.). Cualquier sistema se puede escalar automática o manualmente, de manera que los recursos disponibles parezcan ilimitados.
- Servicio medido: El cliente puede disponer en cualquier momento de un informe de utilización de los recursos. Estos recursos son medidos según un contador que

---

<sup>1</sup> Un navegador, un cliente SSH o la interfaz de comandos de cualquier sistema operativo.

depende del tipo de recurso (horas de ejecución, espacio utilizado, datos transferidos...).

Además de estas características esenciales, la computación en la nube presenta otras ventajas añadidas:

- Costes más bajos
- Facilidad de utilización
- Calidad de servicio
- Fiabilidad
- Administración de IT subcontratada
- Mantenimiento y actualización simplificados
- Barrera (económica) de entrada baja

## 6.2 VIRTUALIZACIÓN

La virtualización es clave en la nube. Permite reunir recursos desde ubicaciones distintas (aunque todas ellas en un centro de datos) y presentarlas al usuario como si se tratase de un único sistema, equiparable a cualquier equipo doméstico.

El hipervisor (Sosinsky, 2012, pág. 142) es el software que permite la virtualización, reúne y gestiona los recursos. Es el programa encargado de “traducir” los recursos reunidos para que el sistema los interprete como hardware independiente. Existen hipervisores de tipo 1, que se ejecutan en el metal desnudo<sup>1</sup> y sobre los cuales se instala el sistema operativo invitado, que será el que utilizará el usuario. Los hipervisores de tipo 2 necesitan un sistema operativo anfitrión, al que el usuario no puede tener acceso.

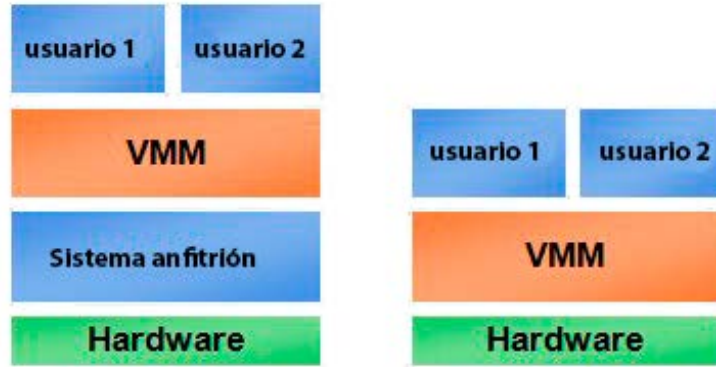
La virtualización implica que los recursos originales nunca podrán ser aprovechados al 100%, dado que existe un software intermedio (el hipervisor) que consume parte de la potencia.

---

<sup>1</sup> “Metal desnudo” es como se conoce al hardware.

Como podemos observar en la imagen a continuación, es la virtualización la que permite la multitenencia. Un mismo hipervisor puede gestionar los recursos de varios usuarios.

*Ilustración 6: datakeeper.com, tipos de virtualización. VMM = hipervisor.  
Traducción de la original libre*



La virtualización además tiene una cualidad importante relacionada con la multitenencia: el aislamiento completo de los distintos clientes. Dos o más usuarios pueden encontrarse utilizando un mismo recurso, y si uno realiza una acción que provoca un fallo grave del sistema, este fallo solo afectará a su instalación, no a aquella con la que comparte recursos.

### 6.3 CENTROS DE DATOS

Un centro de datos es un edificio (o una serie de ellos) en el que se ubican estructuras llamadas “*server racks*”. Son armarios modulares donde se ubican los distintos componentes de la nube.

*Ilustración 7: sala llena de armarios modulares*



Los centros de datos se construyen buscando la eficiencia ecológica. En el caso de Amazon, por ejemplo, la compañía invierte en proveer sus centros de datos de energía solar y eólica (Amazon AWS & Sustainability, s.f.). Un centro de datos suele cumplir las siguientes condiciones (Sosinsky, 2012, pág. 40):

- Acceso a energía de bajo coste.
- Fuentes de energía renovables.
- Cerca de agua abundante (para refrigeración).
- Situados cerca de conexión principal de alta velocidad.
- Costes de terreno modestos.
- Exenciones fiscales.
- Optimizar la latencia<sup>1</sup> global.

### 6.3.1 Economías de alcance

La reducción de costes es fundamental para poder ofrecer a los clientes precios muy competitivos. Amazon ha conseguido ofrecer máquinas virtuales por precios de varios céntimos de dólar por hora de uso. Un centro de datos es un caso típico de economía de alcance (S. Pindyck & L. Rubinfeld, 1998, págs. 197-197):

“Muchas empresas producen más de un producto. A veces estos productos se encuentran estrechamente relacionados entre sí. [...] Para estudiar las ventajas económicas de la producción conjunta, consideremos una compañía de automóviles que produce dos productos: automóviles y tractores. Ambos productos utilizan capital (fábricas y maquinaria) y trabajo como factores. Los automóviles; y los tractores normalmente no se producen en la misma planta, pero sí comparten los recursos de gestión y ambos utilizan una maquinaria y una mano de obra cualificada similares. [...] Generalmente, existen economías de alcance cuando la producción conjunta de una única empresa es mayor que la producción que podrían obtener dos empresas diferentes que produjeran cada una de ellas un único producto (con factores de producción equivalentes distribuidos entre las dos empresas).”

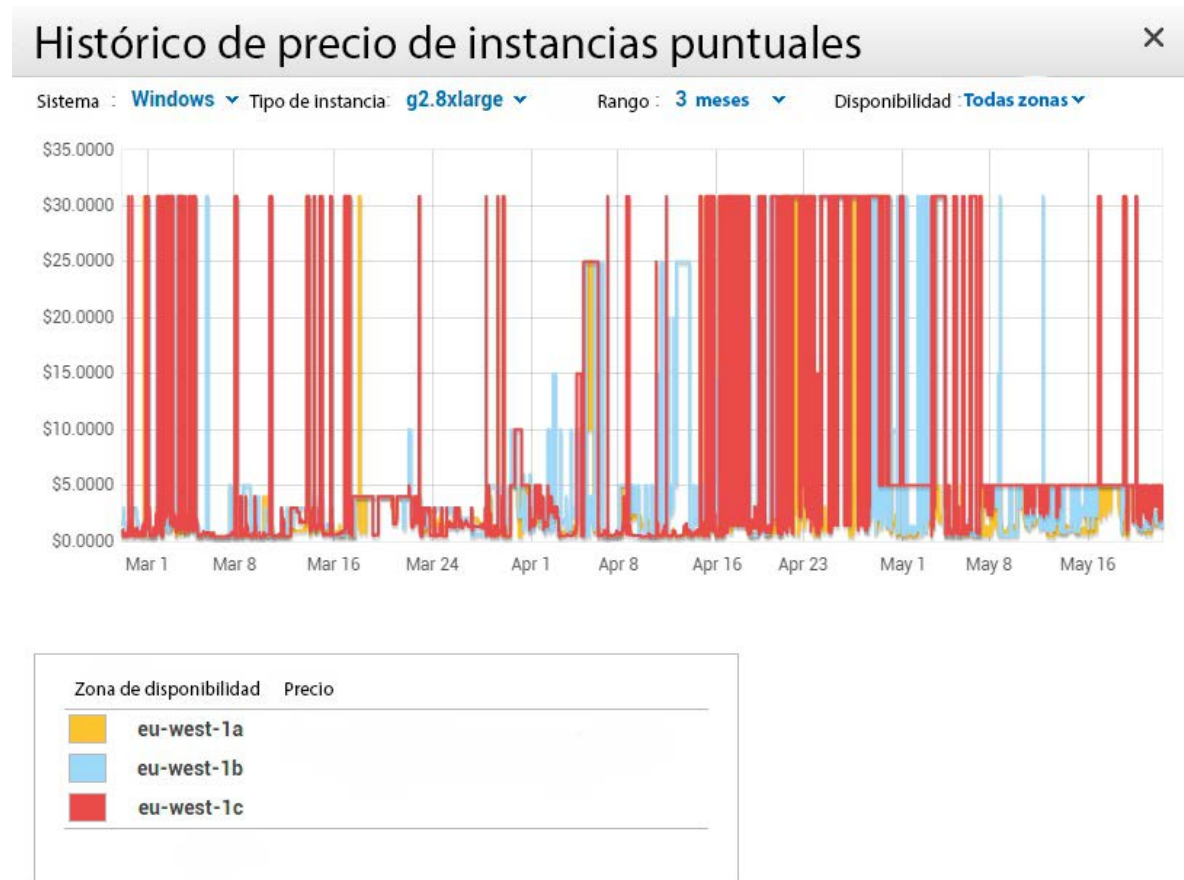
La unión de sistemas de almacenamiento, procesadores, gráficas y otros recursos hace que el aprovechamiento económico de los productos resultantes sea mucho mayor. Además, a esto

---

<sup>1</sup> La latencia, medida en milisegundos, expresa el tiempo que tarda una determinada señal en responder. Por ejemplo, puede medir el tiempo que transcurre desde que escribimos una dirección web y pulsamos intro hasta que el sistema DNS resuelve dicha dirección. La latencia es crucial en sistemas que trabajan a tiempo real, redes sociales, servicios de telecomunicación, etc.

se une la subasta de instancias de AWS (Amazon Spot Instances, s.f.). Se trata de un mercado que permite pujar por la capacidad no utilizada en cada momento, a un precio que fluctúa

*Ilustración 8: fluctuación de precio de las instancias gráficas (g2.8xlarge) a lo largo de 3 meses en el centro de datos de Irlanda. Traducción de la original libre*



según la oferta y la demanda. El usuario elige el precio de su puja (puede verse un gráfico histórico de los distintos precios de instancias) y puede utilizar la capacidad mientras el precio real no supere el que ha establecido. Mediante una reducción del coste de los recursos que de otro modo quedarían sin utilizar se consigue un aprovechamiento máximo de la infraestructura disponible.

De esta manera, una instancia de computación que hace uso de un procesador que en el mercado vale unos 1500 dólares<sup>1</sup> puede ser alquilada por 10 céntimos de dólar cada hora de

<sup>1</sup> Las instancias de computación utilizan procesadores Intel® Xeon® E5-2660 v3

uso. De igual forma, ocupar 1 GB de almacenamiento SSD (gracias al sistema EBS<sup>1</sup>) tiene un precio de 10 céntimos de dólar mensual, cuando el precio de mercado ronda los 30 (ver epígrafe 8.1.1).

### 6.3.2 Regiones y zonas de disponibilidad

Con el objetivo de acercar a los clientes sus servicios (importante para reducir la latencia) Amazon está presente en distintas regiones, dentro de cada cual existen zonas de disponibilidad (Infraestructura global de AWS, s.f.) .

“La nube de AWS opera 33 zonas de disponibilidad en 12 regiones geográficas de todo el mundo, a las que se sumarán a lo largo del año que viene 11 zonas de disponibilidad y 5 regiones más online.”

Las zonas de disponibilidad pueden disponer de uno o varios centros de datos. En Europa, por ejemplo, existen 2 regiones (Irlanda y Frankfurt) con 3 y 2 zonas de disponibilidad respectivamente. En la ilustración anterior pueden verse las 3 zonas de disponibilidad de Irlanda. Hay una tercera región planificada en reino unido.

La transferencia de datos entre zonas de disponibilidad y regiones es gratuita, y además se permite la replicación de los datos. Un uso muy común consiste en establecer máquinas virtuales en cada región que contienen un servidor web. El sistema de DNS de la compañía (otro servicio del paquete AWS) dirige a los visitantes a la máquina virtual más cercana a su posición, de manera que la latencia es siempre la menor posible y la experiencia mejora.

## 7 AMAZON WEB SERVICES

---

### 7.1 AMAZON EC2

Ahora que disponemos de conocimiento básico sobre cómo funciona la nube, empezaremos a describir los servicios en los que se centrará nuestro estudio. El primero de ellos es el servicio de computación de Amazon: *Ec2 (Elastic compute cloud)*.

Ec2 proporciona máquinas virtuales con unos recursos concretos, medidos en capacidad de procesamiento, memoria RAM, almacenamiento, velocidad de acceso a la red y, opcionalmente, capacidad de procesamiento gráfico (con tarjetas NVIDIA). Una instancia (Amazon Ec2 - instancias, s.f.) es una máquina virtual con un sistema operativo limpio y una

---

<sup>1</sup> Elastic block store es un servicio del paquete AWS que proporciona almacenamiento en forma de unidad virtual a EC2. Es diferente al servicio S3, destinado al almacenamiento de “objetos” (archivos simples como imágenes, hojas de estilo, etc.)



serie de recursos asociados. Cada tipo de instancia tiene un precio asociado (Amazon Ec2 - precios, s.f.). En la generación actual existen 7 tipos de instancias:

- T: familia T2, destinadas a uso general. Proporcionan un rendimiento estándar y tienen la capacidad de soportar ráfagas de carga. Estas ráfagas consumen créditos de CPU (1 crédito supone 1 hora con el procesador funcionando al 100%). Dichos créditos se van sumando a un contador mientras la instancia funciona por debajo del 100%. El precio de las instancias T2 varía entre 0,007 y 0,112 céntimos de dólar por hora según los recursos provisionados.
- M: familias M3 y M4, destinadas a uso general. El rendimiento de estas instancias es parecido a la familia T2, pero el sistema de créditos desaparece. El procesador podría funcionar al 100% en todo momento. Su precio oscila entre 0,132 y 0,585 dólares por hora de uso.
- C: familias C3 y C4, optimizadas para computación. Estas instancias están destinadas a hacer un uso intensivo del procesador, y por ello el resto de recursos de la instancia han sido ajustados en consecuencia. Esta familia dispone de opciones que permiten utilizar desde 2 hasta 36 núcleos de procesamiento y entre 3,75 y 60 GB de RAM. El precio según el tipo de instancia va desde 0,119 dólares hasta 1,912 dólares por hora.
- R: familia R3, destinada a un uso intensivo de la memoria. Permite disponer desde 15 hasta 244 GB de RAM. El precio va desde 0,185 hasta 2,964 dólares por hora.
- G: familia G2, de procesamiento gráfico. Estas instancias tienen asignadas entre 1 y 4 tarjetas gráficas NVIDIA con 1536 núcleos CUDA<sup>1</sup> cada una. Optimizadas para tareas gráficas en tiempo real: juegos, representación 3D, etc. Su precio es de 0,702 o 2,808 dólares por hora de uso (solo hay 2 instancias de este tipo).
- I, D: familias I2 y D2, de almacenamiento. Estas instancias priorizan el almacenamiento de datos, y se centran en el rendimiento E/S<sup>2</sup> (instancias I2 con discos SSD) y en la capacidad (instancias D2 con discos HDD). El precio va desde 0,938 hasta 5,88 dólares por hora.

El precio de las instancias que funcionan con sistema operativo Windows es superior en todos los casos, ya que dicho precio incluye un porcentaje destinado a la licencia del sistema operativo.

---

<sup>1</sup> Tecnología propiedad de NVIDIA que permite el procesamiento paralelo desde la tarjeta gráfica. Las GPU actuales son más eficientes en cierto tipo de cálculos que las CPU. Para cálculos financieros o físicos, por ejemplo, es común utilizar GPU como procesador.

<sup>2</sup> Entrada / Salida de operaciones. Cantidad de operaciones de escritura y lectura que ocurren en el disco. Los discos SSD son más eficientes en esta cuestión porque la información se almacena en estado sólido, no en discos que giran constantemente como en el caso de los HDD.

Poner en funcionamiento una instancia de cualquier tipo es muy sencillo y no requiere conocimientos informáticos profundos gracias a la interfaz visual de AWS. Existe una API<sup>1</sup> que permite automatizar algunos procesos, pero está destinada a instalaciones más grandes. En el vídeo adjunto “TFM1” puede observarse el proceso de aprovisionamiento de una instancia en menos de 5 minutos. Este vídeo también está disponible en [https://youtu.be/\\_0pYLVxwF0I](https://youtu.be/_0pYLVxwF0I).

Dado que el renderizado de imagen 3D, vídeo y otros procesos relacionados con la producción se llevan a cabo principalmente en la CPU, las instancias que nos interesan en esta investigación son de tipo C. Más adelante describiremos la estructura y posibilidades de producción, pero como muestra, en el vídeo adjunto “TFM2” se ha realizado una prueba de rendimiento de una instancia tipo C. Dicho vídeo también está disponible en <https://youtu.be/sTpylZqAric>.

Es importante destacar que, a diferencia de lo que ocurre con otros servicios de Amazon, las instancias son independientes en cada región. Esto significa que, aunque pueden estar interconectadas, para que una misma instancia se ejecute en una región distinta debe especificarse explícitamente. Además, en cada región, dependiendo de los costes en los que incurra el mantenimiento del centro de datos, los precios son distintos.

---

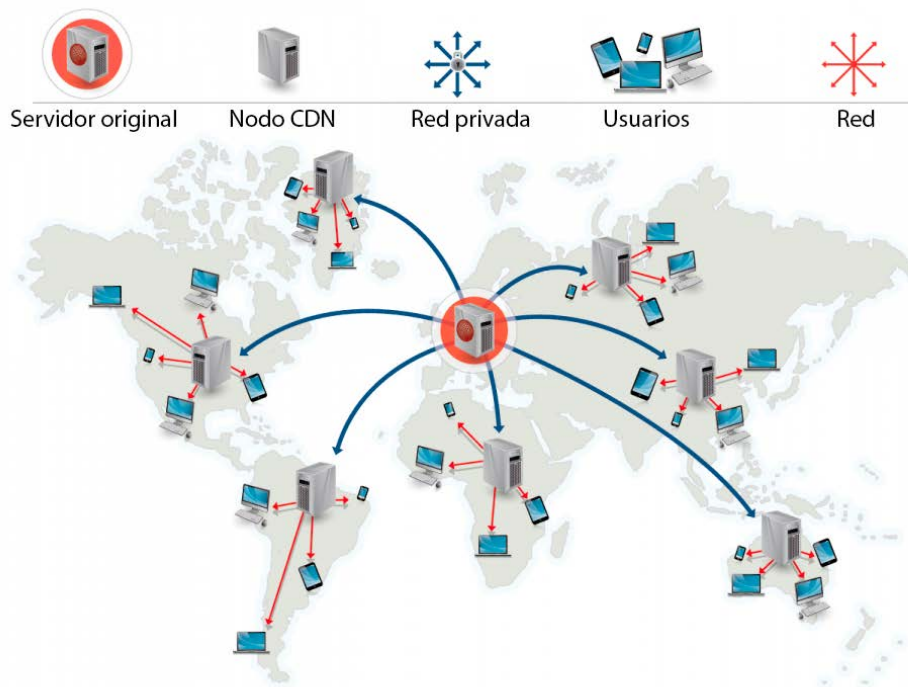
<sup>1</sup> Interfaz de programación de aplicaciones, un conjunto de funciones dispuestas por un proveedor para que cualquier desarrollador pueda programar aplicaciones que se comuniquen con sus servicios. Twitter, Facebook y por supuesto Amazon disponen de API.



## 7.2 CLOUDFRONT

Cloudfront (Amazon Cloudfront, s.f.) es un servicio de CDN (red de entrega de contenidos) distribuido geográficamente. Amazon replica en sus centros de datos de todo el mundo los

*Ilustración 9: funcionamiento esquemático de una red CDN. Traducción de la original libre*



contenidos del usuario, y los asocia a una dirección global. Cuando un visitante accede al contenido, éste es servido desde el centro de datos más próximo, de manera que se reduce significativamente la latencia.

En [https://media.amazonwebservices.com/FS\\_WP\\_AWS\\_CDN\\_CloudFront.pdf](https://media.amazonwebservices.com/FS_WP_AWS_CDN_CloudFront.pdf) puede encontrarse un estudio (en inglés) financiado por Amazon en el que se profundiza en el funcionamiento de Cloudfront y se realizan pruebas. Para lo que nos ocupa, nos basta con saber que, además de mejorar la latencia, alivia la carga de los servidores principales, ya que los objetos más pesados como imágenes, vídeos, etc. se sirven directamente desde los centros de datos de Amazon. Cloudfront comprueba cada cierto tiempo los cambios en los objetos. Si no existen cambios sigue sirviendo la misma copia y si existen, renueva la copia que tiene almacenada.

Los precios de este servicio se reducen según aumenta el uso. Para los primeros 10TB al mes servidos desde Europa, el precio por la transferencia de datos es de 0,085 dólares por GB. A medida que aumenta el uso, el precio por GB disminuye. Europa y EEUU tienen los precios más bajos.

Además de la transferencia, también se facturan las solicitudes (ya sean HTTP o HTTPS). Cloudfront soporta varios tipos de solicitudes. Una solicitud GET por ejemplo sirve para solicitar información a un servidor. Cada vez que accedemos a una página web se realizan cientos de solicitudes GET, una por cada objeto (imagen, hoja de estilos, etc.) que contenga dicha web. Amazon cobra 0,0090 dólares por cada 10.000 solicitudes HTTP y 0,0120 si las solicitudes son HTTPS.

Cada distribución<sup>1</sup> de Cloudfront permite otras opciones como la creación de páginas de error personalizadas o restricciones geográficas, aunque no afectan directamente al rendimiento ni a la latencia.

### 7.3 AMAZON VPC

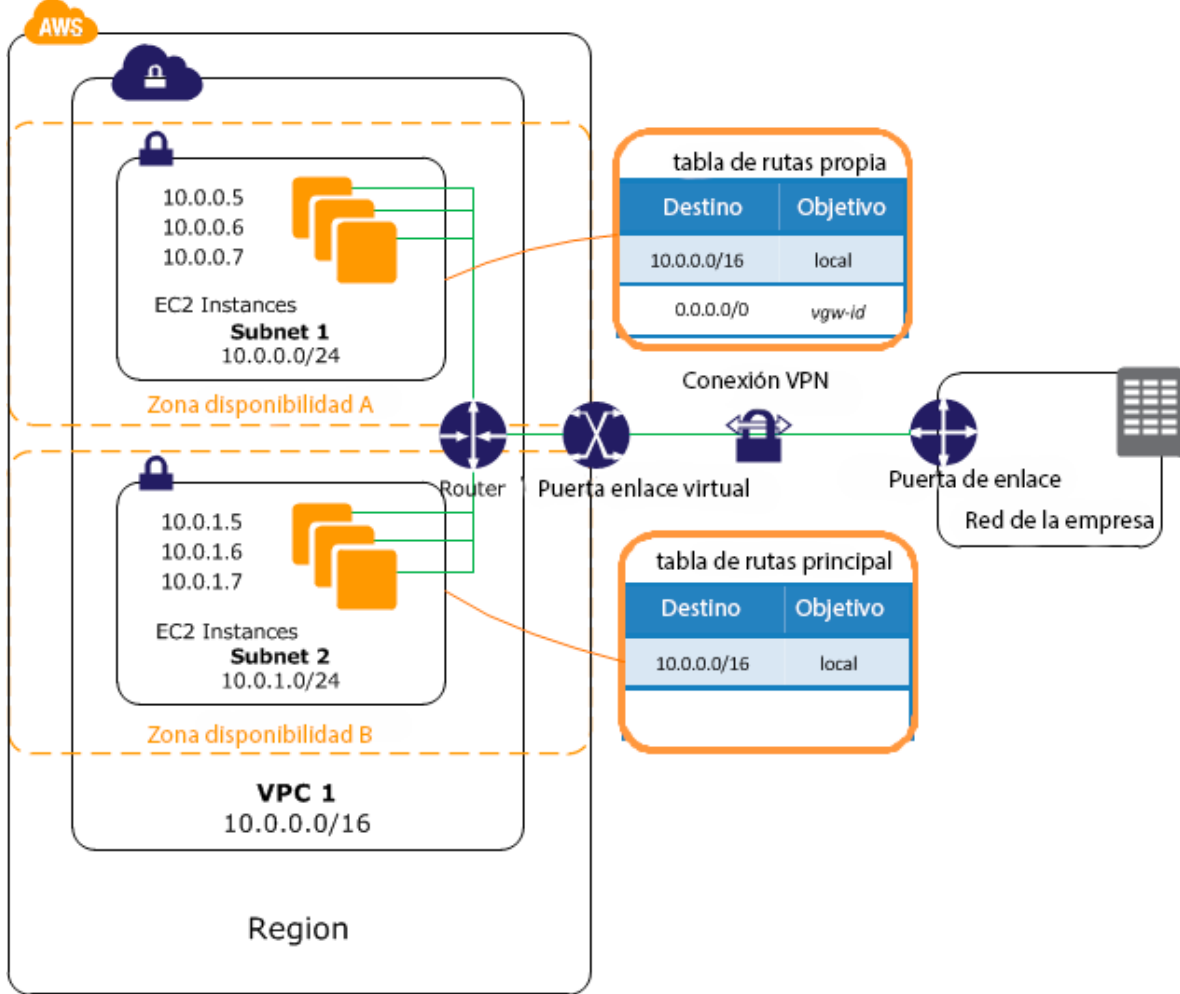
Amazon Virtual Private Cloud (Amazon VPC) es un servicio que permite crear un espacio aislado dentro de la nube de Amazon. Esto añade un nivel más de seguridad a las medidas ya existentes en todos los servicios de AWS. Además, VPC facilita la labor de interconexión de las distintas máquinas virtuales entre sí, lo que será muy útil, por ejemplo, para crear una serie de instancias que trabajen en el renderizado de un mismo recurso. Gracias al servicio NAT Gateway (Amazon VPC, s.f.) es posible establecer conexiones privadas entre instancias y otros servicios AWS, aislándolas del resto de internet.

---

<sup>1</sup> Se conoce como distribución a cada red creada para servir una serie de recursos.

El servicio VPC no se factura tal cual, sino que se establece una tarifa por hora de conexión VPN<sup>1</sup>. Esta conexión VPN cuesta 0,05 dólares por hora y puede utilizarse para conectar cualquier equipo con los servicios de Amazon. Una vez establecida la conexión, los recursos

Ilustración 10: Esquema de funcionamiento de Amazon VPC. Traducción de la original libre



de AWS que están conectados se comportan como si estuviesen dentro de la misma red local. Esta conexión viaja sobre protocolos seguros, de manera que los datos no pueden ser interceptados.

<sup>1</sup> Virtual Private Network es una red virtual privada que, en esencia, funciona como una red local (LAN) pero a través de internet. Los datos que viajan por una VPN lo hacen a través de internet pero son privados.

### 7.3.1 Direct Connect

Amazon Direct Connect es una alternativa para establecer conexión entre el usuario y los servicios AWS. Conectarse mediante VPN implica estar supeditado a la variabilidad de internet, ya que los datos viajan sobre la red pública.

“AWS Direct Connect le permite escalar fácilmente su conexión para adecuarla a sus necesidades. AWS Direct Connect ofrece conexiones de 1 y 10 Gbps, y puede aprovisionar fácilmente varias conexiones si necesita más capacidad. También puede utilizar AWS Direct Connect en lugar de establecer una conexión de VPN a través de Internet con su Amazon VPC, evitando la necesidad de utilizar el hardware de VPN, que con frecuencia no admite velocidades de transferencia de datos superiores a 4 Gbps.”

Con Direct Connect, la conexión es dedicada, aporta mayor fiabilidad y velocidades constantes. El precio aumenta en consecuencia: los precios de la conexión ofrecida directamente por Amazon son de 0,30 y 2,25 dólares por hora de uso. Existe la posibilidad de contratar tarifas menores, pero estas no son ofrecidas directamente por Amazon sino por asociados. El tráfico de Direct Connect viaja también encriptado.

## 7.4 PAGO POR USO

La principal ventaja de los servicios en la nube de Amazon es que ofrecen la posibilidad de pagar a final de mes y solo por el uso. Esto implica que un usuario o empresa puede disponer de equipos por valor de miles de euros en cualquier momento y pagar solo por dos horas de uso, si es eso lo que necesita.

Veamos un ejemplo sencillo. Una instancia c3.8xlarge provee 32 vCPU (es decir, 32 núcleos de procesador), 60 GB de RAM y 2x 320 GB SSD. Estas instancias utilizan procesadores Intel Xeon E5-2680 v2, pero no especifican qué tipo de RAM o SSD. El precio es de 1,912 dólares por hora de uso (sobre Linux, ver epígrafe 8.1.1).

Un procesador de este tipo solo cuenta con 10 núcleos, así que necesitaríamos al menos 4 de ellos (y aun así estaríamos desaprovechando 8 de esos núcleos).

Componente	Cantidad	Precio unitario <sup>1</sup>	Precio total
Procesador	4	1.727 \$	6.908 \$
Memoria (16GB)	4	66 \$	264 \$
Disco SSD (240GB)	3	64,50 \$	193,5 \$
<b>TOTAL</b>			<b>7.365, 5 \$</b>

En la tabla superior se han realizado los cálculos sin tener en cuenta el coste de energía ni otros elementos como placa base, refrigeración, etc. Por el coste de los mismos recursos que nos ofrece AWS, podríamos mantener una instancia a pleno rendimiento durante 3.852 horas, es decir, más de 160 días.

## 7.5 AWS Y RENDERFARMS

Llamamos *renderfarm* o granja de render a un equipo o una serie de ellos preparados específicamente para realizar tareas de procesamiento de vídeo y gráficos 3D. A diferencia de lo que ocurre con una instancia de Amazon o con una instalación propia, en una granja de render no hay que realizar instalaciones previas, en muchas ocasiones, basta con enviar el proyecto y los recursos y recibimos de vuelta el archivo final. Además, los precios por hora de uso suelen ser más bajos que en AWS.

Las principales ventajas frente a AWS son la comodidad y el precio, pero hay inconvenientes que tienen mucho peso. Para empezar, una granja de render no soporta *plugins* o extensiones, o si lo hace, la lista es corta. A día de hoy son muy comunes en cualquier instalación. También es corta la lista de software compatible, de lo que dependerá el formato de proyecto que podamos enviar. En definitiva, se trata de enfrentar comodidad y personalización. Esto puede ser especialmente importante en el caso de software diseñado especialmente para una compañía. En AWS puede instalarse sin problema, en una granja de render no.

## 7.6 ESCALADO Y AUTOESCALADO

Los servicios de Amazon son escalables tanto vertical como horizontalmente (Sosinsky, 2012, pág. 181). Escalar verticalmente supone añadir más recursos a un mismo sistema, por ejemplo, más memoria RAM o capacidad de procesamiento. En el escalado horizontal se añaden “copias” del mismo sistema con recursos idénticos que trabajan en una misma tarea.

---

<sup>1</sup> Precios obtenidos desde el sitio de comercio Amazon.com

El escalado se puede realizar de manera manual, y es lo más común a la hora de escalar verticalmente, sin embargo, también se puede automatizar. El servicio Auto Scaling (Amazon Auto Scaling, s.f.) permite realizar el escalado horizontal de manera automática, estableciendo el número mínimo y máximo de recursos de los que queremos disponer. El propio sistema según unas variables se encarga de decidir cuándo aumenta los recursos y cuándo dejan de ser necesarios y vuelve a reducirlos. Esto provoca una factura mensual más variable, pero asegura que el rendimiento nunca sea causa de fallos.

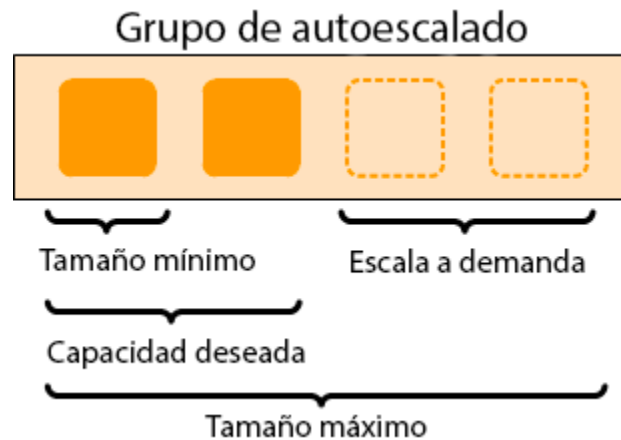


Ilustración 11: esquema de auto escalado. Traducción de la original libre

En el vídeo adjunto titulado “TFM3” se describe el proceso de escalado vertical en AWS. Dicho vídeo también está disponible en [https://youtu.be/2N1wub\\_SAHw](https://youtu.be/2N1wub_SAHw).

## 7.7 OTROS SERVICIOS PARA CONSIDERAR

El paquete de servicios de Amazon es amplio. A continuación, describiremos algunos de los que, si bien no son el centro de este estudio, en determinados casos podrían ser útiles:

- Elastic Transcoder: servicio de codificación de archivos en la nube. Permite utilizar la potencia de codificación de AWS sin necesidad de pasar por máquinas virtuales o instalar software, aunque el precio aumenta. Se factura por fragmentos de 10 minutos de vídeo codificado en SD o HD. Los archivos originales y finales se almacenan en S3.
- S3: Simple Storage Service. Servicio de almacenamiento de objetos estáticos (imágenes, archivos, etc.). Es posible servir páginas web estáticas (basadas en HTML y CSS) desde este servicio, pero no contenidos dinámicos (PHP, etc.). Utilizado en ocasiones como baúl de copias de seguridad.

- Device Farm: pone a disposición del usuario dispositivos móviles reales en los que se puede probar el funcionamiento de aplicaciones. Solo es necesario iniciar la conexión, subir el archivo de la aplicación (.apk en Android) y comenzar a ejecutarla.
- Cloudwatch: servicio de monitorización en la nube. Permite supervisar el correcto funcionamiento de los recursos y enviar alertas en caso de fallos.
- Elastic Load Balancing: servicio de equilibrio de carga. Distribuye el tráfico proveniente de internet entre varias instancias que se están ejecutando. Además, previene errores redirigiendo el tráfico si una de las instancias no funciona correctamente. Puede utilizarse en conjunto con el autoescalado.
- Route 53: servicio de resolución de DNS de Amazon. El sistema DNS “traduce” las direcciones electrónicas legibles ([www.sitioweb.com](http://www.sitioweb.com)) en las direcciones IP originales de los servidores donde están alojados los recursos de dichas direcciones electrónicas.

## 8 PERSPECTIVAS DE MERCADO

El objetivo principal de este documento son las PYME. Una empresa con un tamaño y liquidez suficientes como para no pertenecer a este grupo podría permitirse un gasto inicial en infraestructura que luego podría rentabilizar con el tiempo. Al no poseer la infraestructura ni poder invertir en ella, una PYME se vería relegada a aceptar solo trabajos de baja carga, que pueda permitirse realizar y sobre todo compaginar con otros.

¿Y si eliminamos la principal barrera de entrada? Gracias a AWS, una empresa de cualquier tamaño puede, en cualquier momento, escalar su capacidad de procesamiento hasta niveles teóricamente infinitos pagando solo por el uso, es decir, a un coste de instalación igual a cero.

Vamos a explorar dos posibles contratos que se podrían plantear a una empresa, que se tendría que enfrentar a la necesidad de ampliar su capacidad o rechazar dicho contrato. Analizaremos los costes que supondría el alquiler de capacidad en AWS, su equivalencia en el precio de compra de equipos similares para instalar físicamente y además la estructura de la instalación en la nube resultante.

### 8.1 INDICACIONES PREVIAS

Existen elementos que podrían dificultar las tareas que nos ocupan. El primero y más importante es internet, que supone un cuello de botella importante a la hora de transmitir datos entre los equipos locales y la nube. Para evitar complicar la cuestión, vamos a suponer que las condiciones serán las óptimas para no obstaculizar ni ralentizar el trabajo, aunque es de esperar que en un futuro próximo esto deje de ser un problema, ya que incluso, a día de hoy, hay empresas que ofrecen conexiones asequibles de alto rendimiento (Xataka, 2014).

*Ilustración 12: velocidad de conexión ofrecida por ADAMO.  
Imagen de bandaancha.eu*



Además, Amazon limita el número de instancias de cada tipo que un cliente puede tener en funcionamiento a la vez (Amazon Service Limits, 2014). Esta medida responde a la necesidad de mantener su capacidad como teóricamente infinita y evitar que un solo usuario o empresa cope los recursos de sus centros de datos. No obstante, estos límites pueden cambiarse si se justifica contactando con la empresa. Igualmente, en esta ocasión consideraremos que la



necesidad de computación no superaría los límites y, en caso de superarlos, Amazon autorizaría su ampliación.

### 8.1.1 Tarifas de Amazon Web Services

Las tarifas para cada servicio de Amazon Web Services pueden consultarse en los distintos apartados. Dichas tarifas se han usado con el valor disponible a fecha de febrero de 2016, no obstante, se actualizan constantemente. Se pueden consultar en los siguientes apartados:

- Amazon EC2: <https://aws.amazon.com/es/ec2/pricing/>
- EBS (almacenamiento para instancias: <https://aws.amazon.com/es/ebs/pricing/>
- Transferencia (global): <https://aws.amazon.com/es/ec2/pricing/>
- VPC: <https://aws.amazon.com/es/vpc/pricing/>
- Instancias puntuales: <https://aws.amazon.com/es/ec2/pricing/>
- Route 53: <https://aws.amazon.com/es/route53/pricing/>
- Cloudfront: <https://aws.amazon.com/es/cloudfront/pricing/>
- Cloudwatch: <https://aws.amazon.com/es/cloudwatch/pricing/>

## 8.2 CASO 1: COMPUTACIÓN

### 8.2.1 Situación

Una PYME con un equipo de producción modesto recibe una petición de una agencia publicitaria que requerirá que gran parte del material sea procesado por ordenador. Esta campaña tiene una *deadline* muy cercana (dos semanas) y el equipo que la empresa posee para la creación y renderizado de VFX y 3D no le permite exportar el material con tanta rapidez. Si no quiere rechazar el contrato, la empresa tiene que buscar un método para hacerlo rápidamente mientras continúa trabajando desde el equipo instalado en su oficina.

### 8.2.2 Infraestructura

Ahora que ya sabemos de qué servicios dispone el ecosistema AWS, podemos diseñar la infraestructura que utilizaría la empresa para solucionar su problema.

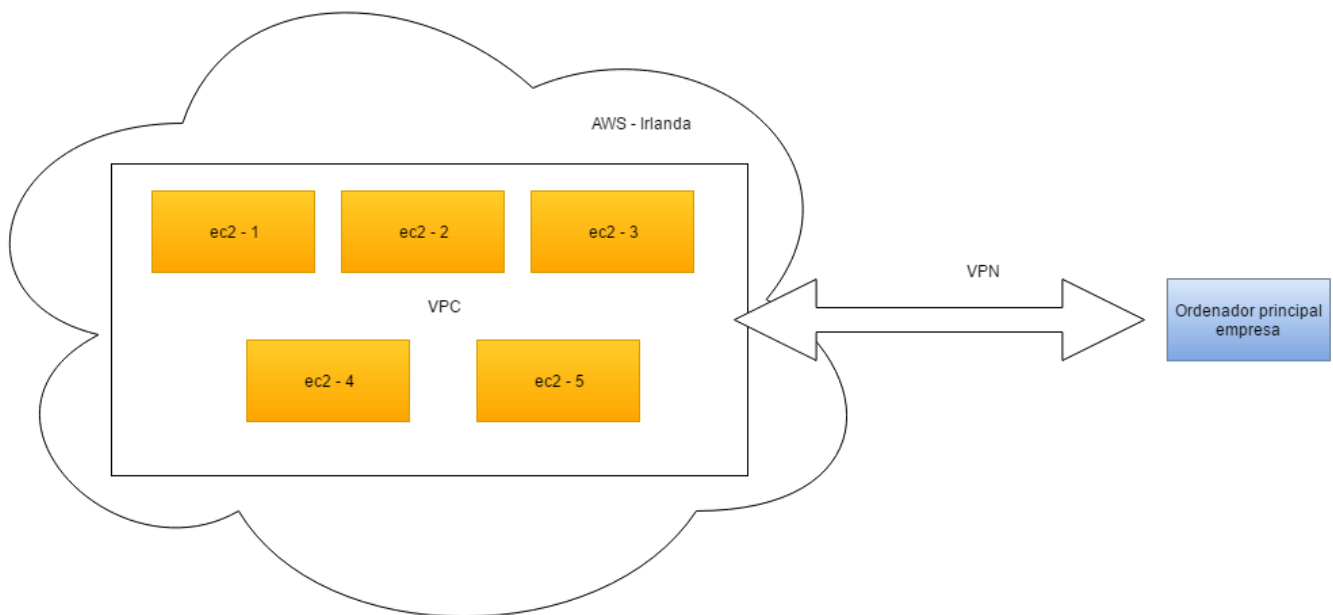
Aunque existen aplicaciones alternativas que pueden funcionar en el entorno de Linux, nos centraremos en aquellas más comerciales que lo hacen sobre Windows. Como ya sabemos, las máquinas virtuales o instancias de Amazon que funcionan con Windows ven incrementado su precio por hora dado que parte de ese precio se destina a pagar la licencia del sistema operativo.

A la hora de renderizar imágenes en 3D, es importante el número de núcleos con los que cuente el procesador. Cinema 4D permite que cada núcleo se dedique al renderizado de una parte de la imagen concreta, con lo cual, a mayor número de núcleos, mayor velocidad de procesamiento. Como ya vimos en el vídeo adjunto TFM2, la instancia de computación más potente permite obtener hasta 36 núcleos que trabajan simultáneamente en una misma imagen.

Además, este software tiene la particularidad de poder trabajar en red. Mediante una herramienta llamada Net Render (Maxon), se puede establecer una red de ordenadores que trabajen en un mismo proyecto.

Supongamos que nuestra empresa, para poder procesar a tiempo todos los elementos 3D, necesitará ejecutar durante las dos semanas 5 instancias de tipo c4.2xlarge (el número máximo permitido según los límites estándar de AWS), lo que proporcionaría un total de 40 núcleos de procesamiento (8 por cada instancia) y 15 GB de RAM para cada una de dichas instancias.

*Ilustración 13: esquema de funcionamiento básico de la infraestructura. Elaboración propia*



Una vez ajustada toda la infraestructura y conectada a la empresa mediante VPN, se podría continuar trabajando en el proyecto mientras, periódicamente, se envía la orden de procesamiento a las máquinas virtuales que se están ejecutando en el centro de datos de Amazon.

En el esquema hemos decidido situar las máquinas virtuales en el centro de datos de Irlanda, que es actualmente el más cercano a España. No obstante, en trabajos de este tipo, donde la latencia no influye en el rendimiento, podría utilizarse un centro de datos de otra región para reducir costes. En el caso de situar nuestra infraestructura de renderización en el centro de datos del norte de Virginia, por ejemplo, el precio se vería reducido de 0,834 dólares por hora a 0,773. Esto supone un ahorro de 0,061 dólares por cada hora de funcionamiento.

Comprobemos cuál sería el coste de mantener esta instalación en funcionamiento durante las dos semanas requeridas para finalizar y entregar el proyecto.

### 8.2.3 Costes

En la instalación anterior hay varios tipos de coste que debemos considerar, los resumimos todos en la tabla siguiente:

Concepto	Cantidad <sup>1</sup>	c/uni <sup>2</sup>	c/total <sup>3</sup>
288h <sup>4</sup> c4.2xlarge	5	0,834	1.200,06
288h VPN	1	0,05	14,4
Almacenamiento <sup>5</sup>	5.000 <sup>6</sup>	0,12	600
Transferencia <sup>7</sup>	3 <sup>8</sup>	0,09	0,27
<b>TOTAL</b>			<b>1.814,73</b>

El coste de mantener esta infraestructura, más que suficiente para las necesidades del proyecto, durante toda una semana, sería de 1.814 dólares, unos 1.600 euros aproximadamente (tasa de conversión 1 de marzo de 2016).

Existe una herramienta dentro de la propia plataforma de Amazon llamada “*TCO calculator*” que permite realizar una comparación del coste de mantenimiento de una infraestructura en las instalaciones y en AWS. Por lo general, el precio que obtendremos será significativamente mayor en el caso de mantener la infraestructura en las instalaciones. No obstante, esta cuestión no es la que nos ocupa.

Como ya comentábamos, la principal ventaja de AWS es que la barrera de entrada es cero, no hay que comprar ni instalar los equipos. Veamos cuánto costaría instalar equipos de características similares en la propia empresa:

---

<sup>1</sup> La unidad dependerá del concepto

<sup>2</sup> Coste unitario en dólares

<sup>3</sup> Coste total en dólares

<sup>4</sup> 14 días x 24 horas de funcionamiento

<sup>5</sup> El almacenamiento SSD para instancias se factura en precio/GB/mes

<sup>6</sup> 1000GB = 1 TB (almacenamiento aproximado)

<sup>7</sup> La transferencia entrante hacia AWS es gratuita

<sup>8</sup> El tamaño total (en GB) del producto final terminado

Concepto	Cantidad	c/uni <sup>1</sup>	c/total
Intel Xeon E5-2666 v3	4	1.445	4.335
Almacenamiento	5x1TB SSD	318	1.590
<b>TOTAL</b>			<b>5.925</b>

El precio de instalar equipos similares en las instalaciones (5.925 dólares) supera de largo lo que pagaría la empresa por alquilar esa capacidad a AWS. Además, en este precio no estamos teniendo en cuenta otro tipo de costes como placas base, memoria RAM (la configuración descrita nos proporcionaba 15 GB por cada instancia), placas base, etc. Además, también habría que cuantificar el coste eléctrico del funcionamiento ininterrumpido de los equipos durante dos semanas.

Es evidente que, en el caso de que la empresa recibiese encargos de suficiente envergadura constantemente, se rentabilizaría de manera rápida la inversión inicial, no obstante, hablamos al principio de unas condiciones en las que este tipo de trabajos sería muy eventual.

#### 8.2.4 Otros usos

Otro de los posibles usos de este tipo de instalación sería la codificación de vídeo. En este caso hay que tener en cuenta la optimización del software utilizado. Como podemos ver en los vídeos adjuntos TFM4 (disponible en <https://youtu.be/LLM31fUb25E>) y TFM 5 (disponible en <https://www.youtube.com/watch?v=VGCBgO5W6rI>) los programas utilizados no son capaces de aprovechar al máximo todos los recursos de las dos instancias de computación más potentes. Habría que encontrar el punto de equilibrio entre el aprovisionamiento de recursos y la máxima capacidad que el software puede utilizar.

Aun así, las instancias disponibles han demostrado que cumplen decentemente con el rendimiento necesario para estas tareas. En el caso de la codificación h.265, que es bastante intensiva (Montes, 2015) el tiempo necesario se reduce casi al 50% en comparación con equipos domésticos y en uno de los perfiles más exigentes (ver gráfico y proceso en el vídeo TFM4).

Estas instancias también pueden utilizarse para tener un escritorio virtual desde el que trabajar a distancia. Podría utilizarse como una estación de trabajo a distancia en la que se estaría editando vídeo desde cualquier dispositivo remoto: un ordenador de sobremesa,

---

<sup>1</sup> Precios obtenidos del portal de comercio Amazon.com

portátil, una tableta... o incluso un explorador web. En términos económicos esta opción no sería tan eficiente, dado que una estación de trabajo está pensada para usarse a largo plazo.

### 8.2.5 Reducción de costes

En la parte descriptiva de este documento hablamos de cómo la subasta de instancias (instancias puntuales o *spot instances*) sirve a Amazon para ocupar al máximo su capacidad disponible a la vez que los clientes ahorran costes. Siguiendo el esquema anterior, pujar por instancias puntuales del mismo tipo tendría el siguiente coste:

Concepto	Cantidad	c/uni	c/total
<b>288h c4.2xlarge</b>	5	0,41 <sup>1</sup>	590,4
<b>288h VPN</b>	1	0,05	14,4
<b>Almacenamiento</b>	5.000	0,12	600
<b>Transferencia</b>	3	0,09	0,27
<b>TOTAL</b>			<b>1.205,07</b>

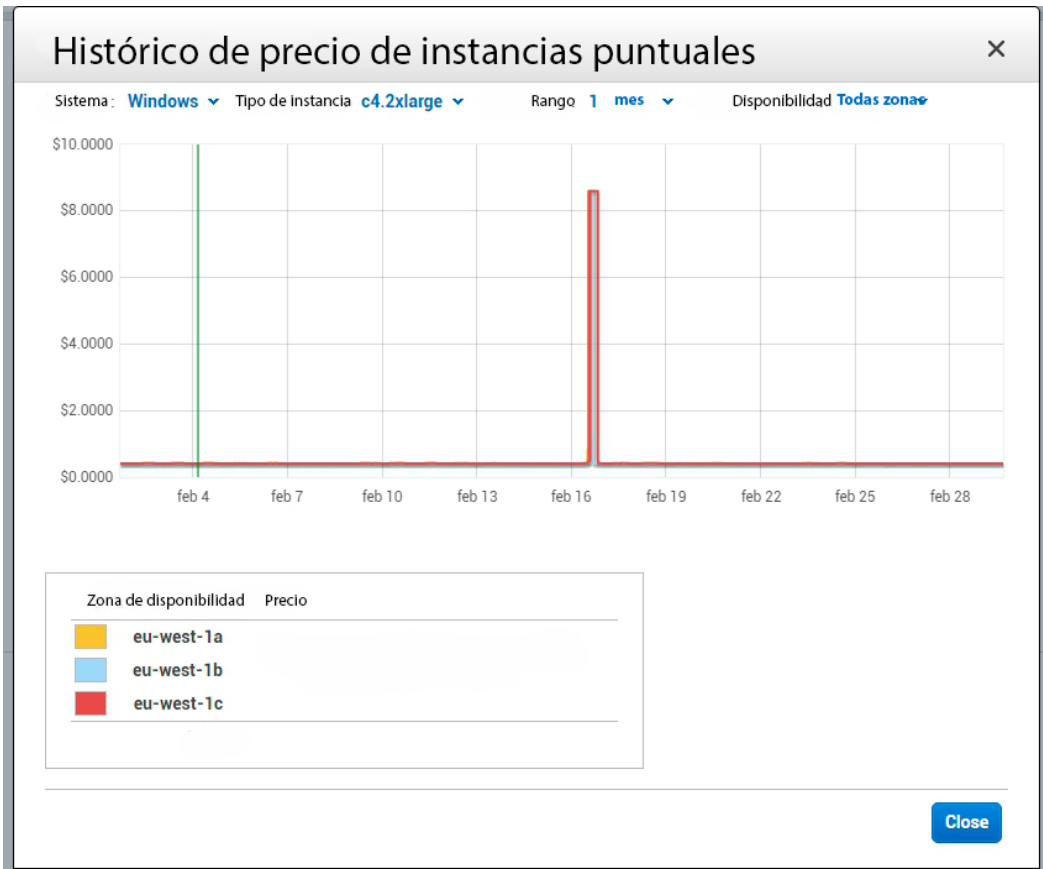
El precio de alquiler utilizando el sistema de subasta se reduce hasta los 1.205 dólares. El total de la reducción viene del coste de funcionamiento de las instancias, dado que el resto de recursos se tarifican igualmente.

---

<sup>1</sup> Precio medio de este tipo de instancias durante los últimos meses.

Las instancias puntuales tienen un riesgo: si el precio de oferta aumenta por encima de la puja, estas se detienen. Sin embargo, la tendencia se puede calcular fácilmente gracias a

*Ilustración 14: histórico de precio de la subasta de este tipo de instancias. Traducción de la original libre*



gráficos históricos disponibles dentro del panel de administración de EC2. En el caso de las instancias c4.2xlarge, este precio es casi invariable, como podemos ver en el siguiente gráfico.

El pico que se observa, probablemente, haya sido debido a una sobrecarga puntual del centro de datos. Esto habría provocado que la capacidad disponible se redujese y, por tanto, aumentase su valor.

Las pujas pueden realizarse por periodos de tiempo concretos, ser indefinidas o persistentes. En los dos primeros casos, la solicitud se mantiene (y la instancia funciona) mientras el precio sea menor al establecido en la puja, y una vez superado, es detenida y eliminada y no se vuelve a ejecutar. En el caso de ser persistentes, si el precio vuelve a bajar, la instancia se ejecuta de nuevo.

## 8.3 CASO 2: WEB INTERACTIVA

### 8.3.1 Situación

La misma empresa tiene que desarrollar una web interactiva que forma parte de la campaña. Se espera una gran afluencia de visitantes, y estos interactuarán bastante en la web (hay un concurso y premios) así que el servidor debe ser lo bastante robusto como para soportar la carga.

### 8.3.2 Infraestructura

Al ser una web que solo estará activa durante una semana, realizar un contrato mensual (lo normal en casi todos los vendedores de espacio web o servidores dedicados) supone un desperdicio de recursos y presupuesto. Además, las visitas serán muy irregulares, experimentando posiblemente una tendencia en forma de campana, con un pico máximo en algún momento de la campaña que acabará decreciendo. Esta tendencia variable haría que, de contratar una cantidad de recursos fija, estos recursos estuviesen desaprovechados durante la mayor parte del tiempo.

Para cumplir con el requerimiento utilizaremos varios servicios de AWS funcionando en conjunto. En primer lugar y como base del sistema provisionaremos instancias de uso general funcionando con Linux<sup>1</sup>. Utilizaremos instancias de uso general, concretamente las instancias m4.large que disponen de dos núcleos de procesador y 8 GB de memoria. Comenzaremos con una instancia general dentro de un grupo de autoescalado, fijando un máximo de 15 instancias, de manera que en los momentos de carga máxima puede haber hasta 30 núcleos y 120 GB de RAM funcionando simultáneamente.

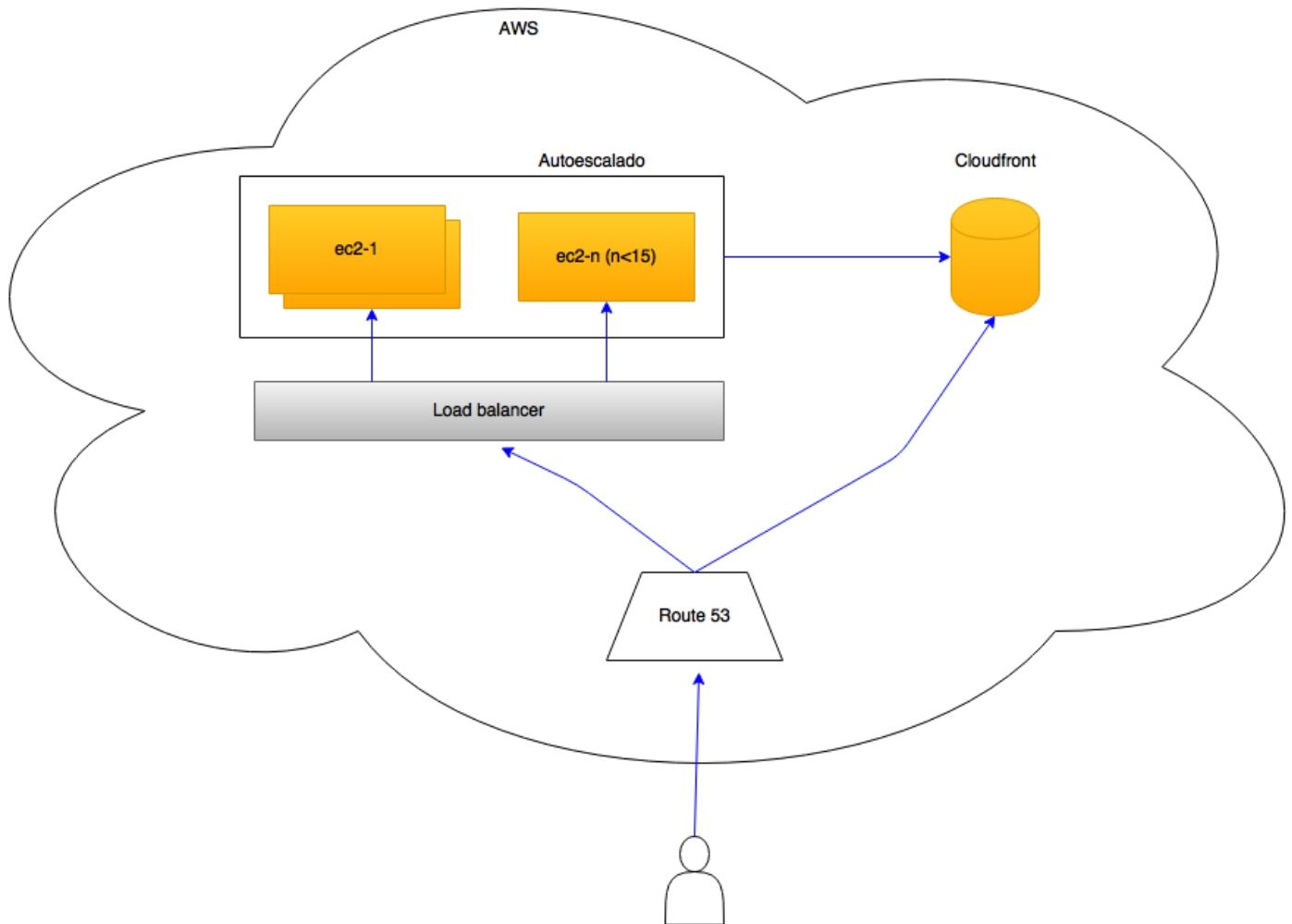
Para acelerar la velocidad de acceso y reducir la carga de los servidores principales, los recursos estáticos (imágenes, hojas de estilo, caché...) se servirán desde Cloudfront. Además, gracias a este servicio, podremos aplicar un certificado SSL para servir los contenidos de manera segura a coste cero, ya que Amazon los ofrece gratuitamente si se utilizan con determinados servicios de su ecosistema.

---

<sup>1</sup> La mayoría de software utilizado en servidores web funciona sobre Linux o alguna de sus variantes

Por último, las DNS serán resueltas por el servicio Route 53 de Amazon. Este servicio, además, permite establecer servidores de respaldo y redirigir a los visitantes a dichos servidores si los principales fallan.

Ilustración 15: esquema de la infraestructura de servidor web con autoescalado. Elaboración propia



### 8.3.3 Costes

El cálculo en este tipo de casos depende de factores muy variables. Para tratar de simplificar las operaciones, supondremos que durante los primeros 2 días solo ha sido necesaria la ejecución de una instancia, los 2 días posteriores el número necesario se eleva a 5 y los 3 últimos días son necesarias 13 máquinas virtuales para soportar el tráfico. Además, la cantidad de usuarios que accederá a la web será de 10 millones, repartidos en 1, 3 y 6 millones respectivamente.



### 8.3.3.1 Periodo 1

Concepto	Cantidad	c/uni	c/total
48h m4.large	1	0,264	12,672
Almacenamiento	8	0,12	0,96
Transferencia <sup>1</sup>	2 x 1mill	0,09	180
Route 53 <sup>2</sup>	1 zona + 1 mill	0,5 + 0,4	0,5 + 0,4=0,9
Cloudfront <sup>3</sup>	1 x 1mill	0,085	85
Cloudwatch <sup>4</sup>	1	3,5/mes	3,5
<b>TOTAL</b>			<b>283,032</b>

### 8.3.3.2 Periodo 2

Concepto	Cantidad	c/uni	c/total
48h m4.large	5	0,264	63,36
Almacenamiento	8 x 5 <sup>5</sup>	0,12	4,8
Transferencia	2 x 3mill	0,09	540
Route 53	1 zona + 3 mill	0,5 + 0,4x3	1,7
Cloudfront	1 x 3mill	0,085	255
Cloudwatch	1	3,5/mes	3,5
<b>TOTAL</b>			<b>868,36</b>

---

<sup>1</sup> El tamaño medio de una página web ronda los 2 MB (Everts, 2015)

<sup>2</sup> Route 53 carga 0,5 dólares por cada zona hospedada y 0,4 por cada millón de peticiones

<sup>3</sup> Supondremos que el peso de los recursos estáticos es de aproximadamente la mitad del peso total de la web, es decir, 1 MB.

<sup>4</sup> El servicio de monitorización Cloudwatch controla el autoescalado.

<sup>5</sup> Aunque se trate de una copia del mismo contenido, cada instancia necesita su propio volumen independiente.

### 8.3.3.3 Periodo 3

Concepto	Cantidad	c/uni	c/total
72h m4.large	13	0,264	274,104
Almacenamiento	8 x 13	0,12	12,48
Transferencia	2 x 6mill	0,09	1.080
Route 53	1 zona + 6 mill	0,5 + 0,4x6	2,9
Cloudfront	1 x 6mill	0,085	510
Cloudwatch	1	3,5/mes	3,5
<b>TOTAL</b>			<b>1.855,984</b>

### 8.3.3.4 Coste total

Periodo	Coste
1	283,032
2	868,36
3	1.855,984
<b>TOTAL</b>	<b>3.007,376</b>

El precio total invertido para soportar la carga de la campaña sería cercano a los 3.000 dólares. Este es un cálculo muy redondeado, ya que las métricas de Amazon son mucho más precisas.

El precio de situar un servidor de características similares en las instalaciones sería el siguiente:

Concepto	Cantidad	c/uni <sup>1</sup>	c/total
Intel Xeon® E5-2676 v3	4	1.589	6.356
Almacenamiento	104 GB SSD	89	89
<b>TOTAL</b>			<b>6.445</b>

Como en el caso anterior, en este presupuesto no estamos contemplando los gastos que supondrían la electricidad y las piezas como memoria RAM, placas base, etc.

Realizar un presupuesto con los precios de los proveedores de servidores dedicados se hace aún más complejo, ya que la mayoría ofrece paquetes cerrados que disponen de una serie de recursos prefijados.

---

<sup>1</sup> Precios obtenidos del portal de comercio Amazon.com

Aún con todo, el objetivo que nos hemos fijado (soportar 10 millones de visitas) es bastante exagerado, estaríamos hablando de que 1/5 de la población de toda España visitaría la web en una semana. Estas cifras podrían responder más bien a una campaña de índole mundial.

Para ser más exactos habría que tener en cuenta que el autoescalado revisa los recursos necesarios cada minuto, y que los recursos de potencia consumidos por una web dependen del número de usuarios conectados simultáneamente, no en total. Por tanto, es bastante probable que durante la mayor parte del tiempo en nuestra campaña, especialmente durante el horario nocturno (lo que sería la mitad) el autoescalado ajustase el número de instancias en funcionamiento a tan solo una.

Como indicábamos al principio, este tipo de cálculos sería demasiado complejo y se escapa al objetivo de esta investigación, por eso, con el objetivo de simplificar, finalmente las cifras son bastante mayores de lo que serían en realidad.

#### **8.3.4 Reducción de costes**

Como ocurría con el caso anterior, se puede hacer uso de las instancias puntuales para reducir el coste operativo. El funcionamiento es el mismo, así que no es necesario volver a describirlo. Teniendo en cuenta el ahorro que obtuvimos en dicho caso, podemos intuir cuál sería en este.

#### **8.3.5 Geocacheado**

Una de las ventajas de AWS y de Cloudfront en particular es que permite el cache geolocalizado. Como Amazon tiene centros de datos en todo el mundo, puede almacenar una copia de los datos en todos ellos, de manera que el acceso a los recursos sea mucho más rápido.

Cuando alguien accede a una dirección web, el servidor DNS más cercano trata de resolver la dirección. Si no la tiene almacenada, la “pregunta” a un servidor DNS central. Una vez obtenida la IP a la que corresponde la dirección, se establece una ruta entre el navegador del usuario y el servidor que, por lo general, pasará por varios nodos, entre ellos un nodo central en el continente de origen y otro en el de destino. Utilizando Cloudfront, una vez resueltas las DNS, la ruta se establece entre el navegador del usuario y el servidor de AWS más cercano que posea una copia de los datos. Todas estas operaciones ocurren en unos milisegundos, pero a la hora de la navegación, unidades tan pequeñas pueden ser las responsables de que un usuario abandone la web desesperado por los tiempos de carga.

### **8.4 RESUMEN**

Los dos casos anteriores se han estudiado de manera esquematizada y simplificada. Elaborar un perfil pormenorizado de las necesidades, casos concretos y sobre todo, presupuestos, requeriría un estudio en profundidad que no entra dentro del objetivo de este estudio.

Basta con decir que las cifras que hemos aportado supondrían una mordida muy pequeña dentro de los presupuestos de una campaña con la envergadura que se solicita. Es perfectamente asumible por una empresa que necesite alquilar esta capacidad de procesamiento o de espacio en web.

La principal diferencia, como ya apuntábamos, es que, gracias a este sistema, una PYME podría ponerse al nivel de cualquier empresa mayor en cuestión de minutos, lo que ayudaría a sus condiciones de competitividad. Si cuentan con los mejores profesionales entre sus filas pero no con la capacidad de procesamiento necesaria para llevar a cabo el proyecto, no hay nada que puedan hacer. Sin embargo, gracias al uso de este sistema, la capacidad dejaría de ser un problema.

Gracias a las herramientas de prueba de rendimiento, cualquier empresa puede realizar test en las distintas instancias de Amazon, en sus máquinas locales, y comparar esos resultados para calcular cuál sería el ahorro en tiempo. Por supuesto, además, podrían seguir utilizando los equipos locales simultáneamente, dado que la carga de procesamiento estaría siendo llevada a cabo en el centro de datos de Amazon.

## **8.5 POSIBLES APLICACIONES ALTERNATIVAS**

Como hemos visto a lo largo del documento, AWS provee gran cantidad de recursos informáticos y estos son teóricamente infinitos. Otra utilidad que puede tener este servicio, además crucial, es la copia de seguridad y la recuperación de desastres. Perder grabaciones es una noticia desastrosa para una empresa audiovisual, ya que en ocasiones estas grabaciones han costado mucho dinero, incluso han podido suponer la destrucción de un set que habría que reconstruir para volver a grabar. Gracias a la capacidad (teóricamente infinita) de almacenamiento de AWS, los datos pueden archivarse en la nube para ser recuperados en caso de pérdida o corrupción de archivos. Estos datos están protegidos mediante conexiones encriptadas durante su transferencia y su acceso solo es posible a usuarios autorizados (además del propio administrador).

## 9 CONCLUSIONES

---

Una de las ventajas del ecosistema de Amazon Web Services radica en su simplicidad. Como hemos podido observar a lo largo de este documento y en los vídeos que lo acompañan, cualquier persona puede crear una cuenta y empezar a experimentar. Realizar tareas sencillas no requiere conocimientos avanzados de informática gracias a su interfaz visual, tan solo cierta familiarización con los conceptos y su funcionamiento.

Los servicios de Amazon están diseñados para trabajar en conjunto, por lo que la optimización es máxima cuando se combinan varios recursos. Resulta más eficiente distribuir los recursos estáticos de una web (mediante Cloudfront) alojados en un servidor (en EC2) y enrutados desde Route 53 que tener cada servicio en un proveedor distinto e interconectarlos.

En el caso de las PYME, la principal ventaja es que no exista barrera económica de entrada al servicio. El material informático de alto rendimiento es caro, y además de ello, la electricidad y la refrigeración requieren unas condiciones y generan unos costes que son difíciles de mantener en el tiempo.

En la conferencia a la que se hace referencia al inicio de este documento (de visionado recomendado), Amazon se propuso compararse con Disney, quien posee más de 55.000 núcleos para renderizar sus películas. La empresa de servicios en la nube consiguió, en su mercado de subasta de instancias, llegar hasta 40.000 núcleos simultáneos. Lo que Amazon demostró con este experimento es que puede llevar una potencia de computación muy similar a la de un gigante como Disney a cualquier empresa del mundo, con el único requisito de disponer de un correo electrónico para abrir una cuenta en AWS y conocimientos informáticos básicos sobre la materia.

En la nube, los datos de cualquier empresa cuentan con la seguridad que aporta una empresa con años de experiencia en el mundo de las comunicaciones y la computación. AWS se toma tan en serio la seguridad que llega a ofrecer certificados SSL gratuitos a sus clientes para que los utilicen con sus servicios. Dichos certificados tienen un valor de entre unos euros mensuales hasta miles. Las conexiones entre los equipos de la empresa y la nube de Amazon se establecen mediante protocolos seguros. Además, para añadir una capa extra, cada usuario puede crear gratuitamente una red privada virtual dentro de la nube, que aísla sus datos y sus recursos de los del resto de clientes.

Otra de las características de la nube es la multitenencia. Gracias a la virtualización, el proveedor de servicios (Amazon en este caso) puede ofrecer a sus clientes recursos compartidos y aislados, de tal manera que dos o más pueden estar utilizando el procesador sin que se interfieran entre ellos ni un error de uno suponga la caída de los demás. Esta característica hace que las instalaciones puedan aprovecharse al máximo y el coste para los clientes sea inferior.

Con servicios como VPC (mediante conexiones VPN) y Direct Connect, una empresa puede crear una infraestructura en la nube y conectarla de manera segura a su red interna y/o a sus servidores, de manera que a efectos prácticos el resultado es el mismo que si dicha infraestructura se encontrase dentro de las instalaciones. Aunque Direct Connect promete una conexión dedicada, en última instancia tanto este servicio con VPN estarán condicionados por la “tubería”, es decir, la infraestructura y las capacidades de la red instalada por los teleoperadores. No obstante es previsible que la situación mejore en los próximos años, como ya indica la “ley del ancho de banda” de Nielsen, que predice que dicho ancho crece un 50% anualmente (Nielsen, 1998). Hasta 2014 estas predicciones se han cumplido, e incluso el crecimiento ha sido mayor de lo esperado.

La creación de una infraestructura dentro de la nube de Amazon es sencilla, como hemos comprobado gracias a los esquemas y los vídeos que acompañan a este documento. Cualquier empresa podría, en un momento determinado en el que su producción lo requiera, aprovisionar recursos del fondo común de AWS y utilizarlos durante un tiempo concreto, creando una especie de infraestructura desechable que no incurre en gastos de montaje ni desmontaje, solo de uso u operacionales. Además los recursos pueden escalarse automáticamente o manualmente.

Siempre desde el punto de vista de un uso puntual, utilizar la infraestructura de AWS supone un ahorro con respecto a invertir en una infraestructura en las instalaciones. Es obvio que, para un uso continuado, una inversión en material acaba rentabilizándose, tarde o temprano. Sin embargo, cuando se trata de un encargo esporádico, el hecho de poseer o no esa infraestructura supone aceptar o rechazar el trabajo. Mantener dicha infraestructura en las instalaciones a la espera de posibles trabajos eventuales supondría un lastre económico para la empresa, sin embargo, gracias a AWS, es posible almacenar una “imagen” de dicha infraestructura para mantenerla inactiva y, en caso de necesidad, poder recuperarla en cuestión de minutos.

La presencia internacional de los centros de datos de Amazon hace posible la optimización de la entrega de contenidos desde el punto más cercano al usuario. Esta característica no solo reduce la carga de los servidores centrales de una página web o aplicación, por ejemplo, sino que disminuye el tiempo de acceso a los recursos. Este tiempo de acceso, que puede ser de apenas unos milisegundos, supondría la diferencia entre fidelizar a un usuario o perderlo para siempre. Tenemos por ejemplo el caso de las caídas del servicio de Whatsapp y la migración masiva de usuarios a aplicaciones alternativas como Telegram. Una caída del servicio es algo que aporta una mala experiencia al usuario, y puede evitarse muy fácilmente.

Amazon Web Services no es el Santo grial de los servicios en la nube. Muchos usuarios lo consideran uno de los proveedores menos económicos del sector, y además su infinidad teórica de recursos se desmonta a poco que se investigan los límites del servicio. Esto último es lógico, teniendo en cuenta que lo infinito tan solo existe en el cálculo matemático. En la

realidad, los recursos que AWS puede ofrecer estarán limitados por la capacidad del centro de datos. No obstante, si por algo destaca, y es por lo que ha sido el centro de este documento, es por su facilidad de acceso. Desde que se pusiese a disposición del público su interfaz gráfica, solicitar recursos ha estado al alcance de cualquier persona. En cinco minutos y varios clics es posible tener una infraestructura bajo demanda funcionando en un centro de datos de Irlanda, por ejemplo, cuyo coste de entrada es cero y que ofrece un repertorio de tarifas muy competitivas.

Toda la investigación se ha centrado en la rentabilidad de la nube frente a las instalaciones físicas dentro de una empresa. El punto de rentabilidad debe estudiarse en cada caso concreto, según la empresa y su ritmo de trabajo. A partir de cierta cantidad de “trabajos puntuales” que requieren mayor capacidad, el coste de alquiler en AWS superaría el generado por la compra y mantenimiento de equipos dedicados. Según los presupuestos expuestos, apenas tres trabajos de este tipo (al año) ya serían suficientes para que el coste de instalación (no de mantenimiento) resultase inferior al de alquilar infraestructura en esas 3 ocasiones. Como indicábamos, debería estudiarse en cada caso de empresa concreta, ya que estaría sujeto a la variabilidad del mercado.

## 10 RECURSOS AUDIOVISUALES

---

Para facilitar la comprensión de algunos aspectos de esta investigación se han creado vídeos que muestran los servicios de AWS y su funcionamiento. Estos vídeos (que se adjuntan en formato físico) están disponibles en la red (a fecha de junio de 2016) para su visualización y descarga:

- TFM1: Iniciar máquina virtual en AWS, disponible para visualización y descarga en <http://tfm-manmon01.s3.amazonaws.com/hls/TFM1/index.html>
- TFM2: Prueba de rendimiento de instancia de computación, disponible para visualización y descarga en <http://tfm-manmon01.s3.amazonaws.com/hls/tfm2/index.html>
- TFM3: Escalar verticalmente, disponible para visualización y descarga en <http://tfm-manmon01.s3.amazonaws.com/hls/tfm3-escalar/index.html>
- TFM4: Rendimiento y codificación 8xlarge, disponible para visualización y descarga en <http://tfm-manmon01.s3.amazonaws.com/hls/tfm4-codif/index.html>
- TFM5: Rendimiento y codificación 4xlarge, disponible para visualización y descarga en <http://tfm-manmon01.s3.amazonaws.com/hls/tfm5-codif4xlarge/index.html>

## 11 BIBLIOGRAFÍA

---

- Hernández Sampieri , R., Fernández Collado, C., & Baptista Lucio , P. (1991). *Metodología de la investigación*. México: McGRAW - HILL INTERAMERICANA DE MÉXICO.
- S. Pindyck, R., & L. Rubinfeld, D. (1998). *Microeconomía* (Cuarta ed.). Madrid: Prentice Hall.
- Sosinsky, B. (2012). *¿Qué es la nube? El futuro de los sistemas de información*. (M. Pascual Cabrerizo, Trad.) Madrid: Anaya Multimedia.

## 12 CITAS DE INTERNET

---

- Amazon. (s.f.). *About us: AWS*. Recuperado el 5 de diciembre de 2015, de <http://aws.amazon.com/about-aws/>
- Amazon Auto Scaling. (s.f.). *User guide*. Recuperado el 20 de febrero de 2016, de <http://docs.aws.amazon.com/autoscaling/latest/userguide/WhatIsAutoScaling.html>
- Amazon AWS & Sustainability. (s.f.). *AWS & Sustainability*. Recuperado el 22 de enero de 2016, de AWS & Sustainability: [https://aws.amazon.com/es/about-aws/sustainability/?nc1=h\\_ls](https://aws.amazon.com/es/about-aws/sustainability/?nc1=h_ls)
- Amazon case Studies. (s.f.). *Casos prácticos e historias de éxito de clientes impulsados por la nube de AWS*. Recuperado el 3 de febrero de 2016, de Amazon Web Services: <https://aws.amazon.com/es/solutions/case-studies/>
- Amazon Cloudfront. (s.f.). *Descripción de Cloudfront*. Recuperado el 10 de febrero de 2016, de <https://aws.amazon.com/es/cloudfront/>
- Amazon Ec2 - instancias. (s.f.). *Tipos de instancias*. Recuperado el 18 de enero de 2015, de <https://aws.amazon.com/es/ec2/instance-types/>
- Amazon Ec2 - precios. (s.f.). *Precios de Ec2*. Recuperado el 17 de febrero de 2016, de <https://aws.amazon.com/es/ec2/pricing/>
- Amazon EC2. (s.f.). *Amazon EC2, descripción*. Recuperado el 10 de enero de 2016, de <https://aws.amazon.com/es/ec2/>
- Amazon EC2 reserved instances. (s.f.). *Instancias reservadas de Amazon EC2*. Recuperado el 15 de marzo de 2016, de <https://aws.amazon.com/es/ec2/purchasing-options/reserved-instances/>



- Amazon Service Limits. (19 de junio de 2014). *Amazon EC2 Service Limits Report Now Available*. Recuperado el 20 de diciembre de 2015, de <https://aws.amazon.com/es/about-aws/whats-new/2014/06/19/amazon-ec2-service-limits-report-now-available/>
- Amazon Spot Instances. (s.f.). *Instancias de subasta de Amazon*. Recuperado el 14 de febrero de 2016, de <https://aws.amazon.com/es/ec2/spot/>
- Amazon VPC. (s.f.). *NAT gateways*. Recuperado el 16 de febrero de 2016, de <http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/vpc-nat-gateway.html>
- AWS re:Invent. (2015). (CMP404) Cloud Rendering at Walt Disney Animation Studios. *AWS re:Invent 2015*. Recuperado el 1 de noviembre de 2015, de <https://youtu.be/btGphZZqkY>
- Axceleon Inc. (2010). Recuperado el 13 de diciembre de 2015, de [http://www.cloudfuzion.com/aws/CloudFuzion\\_EC2\\_White\\_Paper.pdf](http://www.cloudfuzion.com/aws/CloudFuzion_EC2_White_Paper.pdf)
- Azure CDN. (s.f.). *Microsoft Azure CDN*. Recuperado el 15 de marzo de 2016, de <https://azure.microsoft.com/es-es/services/cdn/>
- Cid, M. (11 de marzo de 2014). Adamo rompe en velocidad y precio: 1000 Mbps de bajada y 200 Mbps por 38.14 euros al mes. *XatakaON*. Recuperado el 21 de febrero de 2016, de <http://www.xatakaon.com/tarifas-adsl-y-cable/adamo-rompe-en-velocidad-y-precio-1000-mbps-de-bajada-y-200-mbps-por-38-14-euros-al-mes>
- Dirección General de Industria y de la PYME. (2015). *RETRATO DE LAS PYME 2015*. Recuperado el 22 de diciembre de 2015, de [http://www.ipyme.org/publicaciones/retrato\\_pyme\\_2015.pdf](http://www.ipyme.org/publicaciones/retrato_pyme_2015.pdf)
- Everts, T. (11 de junio de 2015). *Page bloat update: The average web page is more than 2 MB in size*. Recuperado el 25 de febrero de 2016, de Soasta.com: <https://www.soasta.com/blog/page-bloat-average-web-page-2-mb/>
- Google CDN. (s.f.). *Google Cloud CDN*. Recuperado el 15 de marzo de 2016, de <https://cloud.google.com/cdn/docs/>
- Google Cloud Preemptible VM. (s.f.). Recuperado el 15 de marzo de 2016, de <https://cloud.google.com/preemptible-vms/>
- Google Compute Engine. (s.f.). *Google Compute Engine Pricing*. Recuperado el 15 de marzo de 2016, de <https://cloud.google.com/compute/pricing>
- Greene, J. (19 de abril de 2015). *Sky-high expectations for release of Amazon Web Services' numbers*. Recuperado el 20 de marzo de 2016, de The Seattle Times:

<http://www.seattletimes.com/business/amazon/sky-high-expectations-for-release-of-amazon-web-services-numbers/>

Infraestructura global de AWS. (s.f.). AWS. Recuperado el 2015 de noviembre de 20, de <https://aws.amazon.com/es/about-aws/global-infrastructure/>

Kennedy, J., & Healy, P. (2014). *Estados Unidos Patente n° US20140237373 A1*. Recuperado el 27 de diciembre de 2015, de <https://www.google.com/patents/US20140237373>

Lynley, M. (28 de abril de 2016). *Amazon's Web Services are shining in its latest earnings beat*. Recuperado el 17 de marzo de 2016, de Techcrunch.com: <http://techcrunch.com/2016/04/28/amazon-is-spiking-after-posting-a-huge-earnings-beat/>

Maxon. (s.f.). *Net Render*. Recuperado el 24 de enero de 2016, de cinema4d.com: <http://www.cinema4d.com.mx/modulos/net-render.html>

maxon.net. (s.f.). *Cinebench Overview*. Recuperado el 5 de febrero de 2016, de <http://www.maxon.net/products/cinebench/overview.html>

Montes, M. (22 de noviembre de 2015). *El códec H.265: impresiones*. Recuperado el 15 de abril de 2016, de Manuel5cc.es: <http://www.manuel5cc.es/comunicacion-audiovisual/el-codec-h-265-impresiones/>

Nielsen, J. (5 de abril de 1998). *Nielsen's Law of Internet Bandwidth*. Recuperado el 10 de marzo de 2016, de Nielsen Norman Group: <https://www.nngroup.com/articles/law-of-bandwidth/>

Staveley, M. (s.f.). *Applications that scale using GPU Compute*. Recuperado el 15 de marzo de 2016, de <https://azure.microsoft.com/es-es/documentation/videos/azurecon-2015-applications-that-scale-using-gpu-compute/>

Xataka. (11 de marzo de 2014). *Adamo rompe en velocidad y precio: 1000 Mbps de bajada y 200 Mbps por 38.14 euros al mes*. Recuperado el 15 de diciembre de 2015, de <http://www.xatakaon.com/tarifas-adsl-y-cable/adamo-rompe-en-velocidad-y-precio-1000-mbps-de-bajada-y-200-mbps-por-38-14-euros-al-mes>